Biological Safety Factors in Toxicological Risk Assessment

Health and Welfare Canada
Santé et Bien-être social Canada
Canada

90-06409/

90-EHD-154

Biological Safety Factors in Toxicological Risk Assessment

A report produced for the:

Environmental Health Directorate
Health Protection Branch
Health and Welfare Canada

by
R. Stephen McColl, Ph.D.
Department of Health Studies
University of Waterloo
and the
Institute for Risk Research
University of Waterloo

Également disponible en français sous le titre:
«Facteurs de sécurité biologiques et évaluation du risque toxicologique»

Copies of this publication may be obtained from:
Communications Branch
Department of National Health and Welfare
5th Floor
Brooke Claxton Building
Ottawa, Ontario
K1A 0K9

©Minister of Supply and Services Canada
Cat. H49-49/1990E
ISBN 0-662-17638-3

This report was prepared under a research contract from the Environmental Health Directorate, Health Protection Branch, Health and Welfare Canada.

The author of the report is R. Stephen McColl, Ph.D., who is a member of the Institute for Risk Research, and a faculty member in the Department of Health Studies, University of Waterloo.

Dr. D. Krewski and Dr. V. Armstrong of the Environmental Health Directorate, Department of National Health and Welfare, reviewed the report and provided technical assistance.

The views and recommendations contained herein represent the opinion of the contractor and do not necessarily constitute endorsement by the Department of National Health and Welfare. Data, models, calculations, and recommendations presented in this study are for illustrative purposes only.

March, 1989.

Table of Contents

		Page
1.	**Fundamental Concepts in Safety Factor Determination**	**7**
A.	The Threshold Concept	7
B.	Classification of Toxic Effects	8
2.	**Acceptable Daily Intake and Safety Factors**	**10**
A.	Development of the Safety Factor Concept	10
B.	NOEL Safety Factor Methodology	12
	1. Definition of ADI	12
	2. Definition of NOEL and Other Threshold Estimators	12
C.	Specification of Safety Factors	14
	1. Definition of Safety Factor	14
	2. Selection of Safety Factors	15
D.	Validation of Safety Factors	16
	1. Overall Safety Factor	16
	2. Specific Safety Factors	16
	a) Intraspecies Variation	16
	b) Interspecies Extrapolation	17
	c) Subchronic to Chronic Adjustment	17
	d) LOEL to NOEL Adjustment	17
3.	**Alternative Approaches to Safety Factor Methods**	**19**
A.	Criticisms of the NOEL-SF Method	19
B.	Mathematical Safety Factor Models	20
	1. Normal Distribution (Probit) Models	20
	2. Gaylor Linear Extrapolation Method	20
	3. Dourson Safety Factor Method	22
	4. Crump Benchmark Dose Method	22
	5. Summary and Critique	23
C.	Severity of Effects	24
	1. Severity Scales	24
	2. NOEL Estimated by Dose-Severity Diagrams	25
4.	**Safety Factor Approaches for Carcinogens**	**27**
A.	General Background	27
B.	Traditional Safety Factors for Carcinogens	27
C.	Weight-of-Evidence Classification Approach	28
D.	Mechanistic Classification Approach	29
E.	Mechanistic Approaches for the Risk Assessment of Dioxin	31
	1. EPA Methodology	31
	2. OME Methodology	32
F.	PB-PK Modeling	33
5.	**Safety Factor Approaches for Teratogens**	**35**
A.	General Background	35
B.	Noel-SF Approach for Teratogens	36
C.	Mathematical Modeling Approaches for Teratogens	37
D.	Interspecies Extrapolation	38
E.	Special Problems of Safety Factor Determination for Teratogens	38
F.	Summary of Safety Factor Determination for Teratogens	38
	References	**40**

1. Fundamental Concepts in Safety Factor Determination

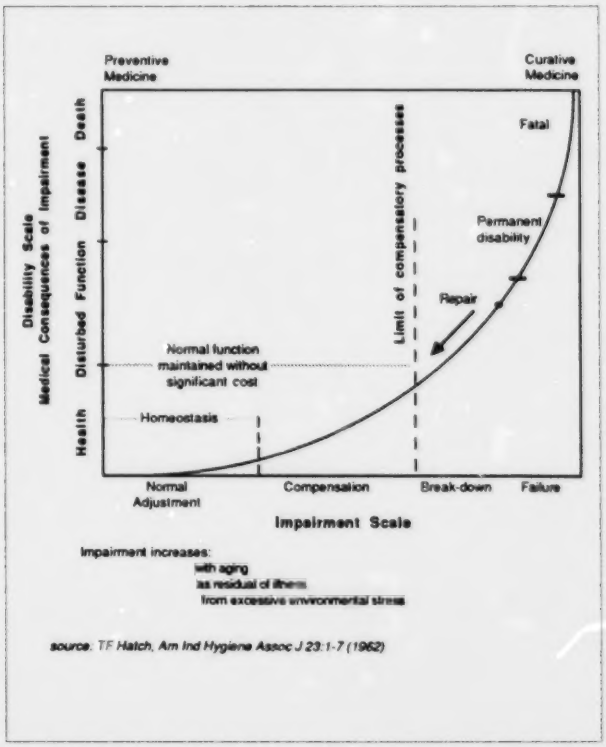
A. The Threshold Concept

Traditional approaches to toxicity testing for environmental hazards have focused primarily on the determination of "safe" levels of exposure to toxic agents. Within this conceptual framework, the definition of "safety" is equated with the absence of toxic effects in the exposed human population. Both in theory and in practice, however, toxicological studies can only provide estimates of minimum levels of risk as operationally defined by appropriate animal tests or epidemiological studies, rather than predicting levels of absolute safety (Sharratt, 1976; Kroes and Feron, 1984; Rodricks and Tardiff, 1984).

The estimation of safe levels of exposure for most toxicants is based on the *threshold concept*, which states that the dose level determines whether or not a pathological effect will occur upon exposure. Conventional toxicological theory postulates that small doses of a toxic agent can be tolerated due to the presence of systems for metabolic detoxication, physiological homeostasis, and cellular adaptation and repair. Below a certain minimum dosage, these compensatory mechanisms can minimize the effects of a toxic agent, and therefore a toxicological threshold may be defined as a dose below which no pathological changes to the organism will occur (Aldridge, 1986). At higher dose levels, however, the compensatory ability of body defense systems becomes saturated, leading to a loss of physiological homeostasis and an impairment in organ function (Figure 1.1). Eventually physiological impairment manifests itself as one or more pathological processes that lead to a disease state in the exposed organism. This marks the threshold for a particular *critical endpoint*.

In practice, the process of determining toxicity thresholds is actually more complicated, since with increasing dosage there will be a series of toxic events in several target organs, as different defensive or compensatory mechanisms are overwhelmed at varying dose levels (Hattis, 1987). Therefore several different thresholds may have to be defined according to multiple endpoints of primary concern for human health. An *individual threshold* must be determined for each critical endpoint, defined as a subthreshold exposure level insufficient to produce the toxic effect within a single person. Moreover, a *population threshold* must also be defined according to the cumulative response of all individuals to exposure to the toxic agent, taking into account the natural variation in sensitivity that exists within the population.

Figure 1.1
The Impairment-Disability Curve



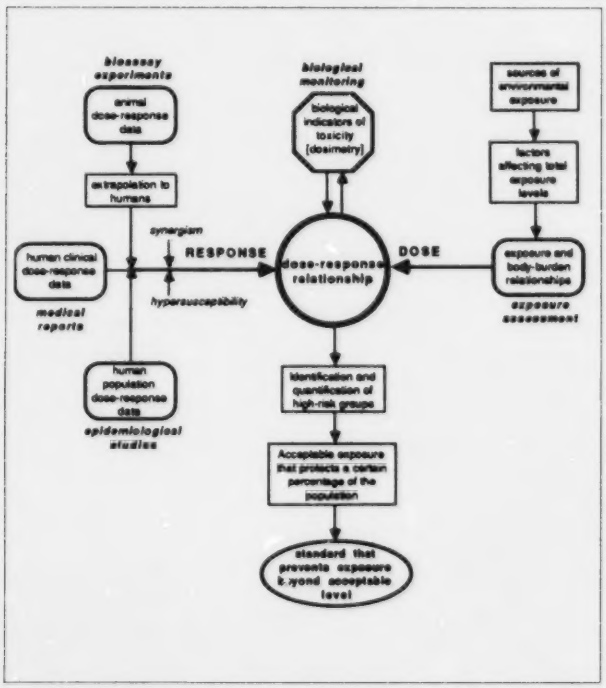
Given the lack of human toxicity data for many hazardous substances, there is often considerable uncertainty about the distribution of toxic thresholds within a population, especially for the most sensitive individuals. In addition, multiple exposures to mixtures of toxic chemicals in the environment may produce unpredictable synergistic or potentiation effects, which can significantly increase the toxicity of chemicals (Figure 1.2). Finally, since laboratory animals are frequently

used as surrogates for exposed human populations, any threshold values experimentally derived from animal tests must be transformed into equivalent human exposure measures, using interspecies extrapolation techniques whose accuracy is affected by the inclusion of several assumptions and uncertainties.

B. Classification of Toxic Effects

Environmental health hazards do not invariably act by the same basic type of pathological mechanism. Thus the selection of a particular risk assessment methodology will depend on the mechanistic classification of the toxicity effect. A useful categorization scheme has been described by Hattis (1987), which divides toxic effects into two fundamental classes according to mechanisms of disease causation and reversibility of effects (Table 1.1).

Figure 1.2
Model for Deriving Exposure Standards



"Traditional" toxicity assessment is based on a *deterministic* mechanism of toxicity. Deterministic processes act by overwhelming the body's compensatory processes by exposure to a sufficiently high dose of a toxic agent, thereby exceeding the toxicity threshold and leading to pathological effects. These effects are, however, reversible until the later stages of the disease process (Figure 1.1). This category is generally referred to as *systemic toxicity*, and includes both traditional acute toxicity and traditional chronic toxicity for any health effects not associated with cancer or gene mutation (Barnes and Dourson, 1988). The pathological mechanisms giving rise to systemic toxicity can be summarized as follows:

> *When a biological function that depends on the collaborative action of a large number of the same macromolecules such as receptors or enzymes is deranged then it seems probable that a threshold exists. The action of a number of molecules of a receptor or enzyme could be prevented or modified without any demonstrable change in biological function. However, when the biological function is dependent on the action of one molecule such as DNA or those molecules (enzymes) associated with it in a one-to-one relationship, then the existence of a threshold seems fundamentally unlikely. Acceptance of this view, of course, means that exposure to most toxic chemicals at low concentrations for a lifetime will be harmless. In contrast, for those chemicals causing genetic changes this seems unlikely. Such statements ignore questions of risk involving the slope of the dose response curves ... or the differing susceptibility of individuals. Practical decisions have to be made even though the risk is not zero.*
>
> Aldridge (1986)

Table 1.1
Types of Health Hazards Requiring Fundamentally Different Risk Assessment Approaches

"Traditional" Deterministic Toxicity Proceeds by overwhelming body compensatory processes; below some threshold, in individuals who are not already beyond the limits of normal function without exposure, response is reversible.
• Traditional Acute Toxicity Toxic damage is completely reversible or proceeds to long-term damage within about three days of exposure, e.g., lung damage from inhalation of chlorine.
• Traditional Chronic Toxicity Toxic process typically proceeds to permanent damage over a time period from several days to several months, due to either (A) reversible accumulation of a toxic agent (e.g., methyl mercury) or (B) accumulation of a slowly reversible toxic response.
Non-Deterministic Toxicity Effects resulting from insidious processes that are irreversible or poorly reversible at low doses or early stages of causation.
• Molecular Biological (Stochastic Process) Effects Occur as a result of one or a small number of irreversible changes in information coded in DNA, e.g., mutagenesis, most carcinogenesis, and some teratogenesis.
• Chronic Cumulative Effects Occur as a result of a chronic accumulation of many small-scale damage events, e.g., birth defects, emphysema, noise-induced hearing loss, atherosclerosis, and probably hypertension.

Adapted from Hattis, 1987.

Given the assumption of a threshold mechanism, the traditional approach to assessing systemic toxicants has been based on the concept of *acceptable daily intake* (ADI) — a toxicological parameter defined as the amount of a substance that can be taken into the body on a daily basis over a prolonged period of time without suffering significant adverse health effects. A *safety factor* (SF) is usually included in the calculation of the ADI value to account for the uncertainties in estimating the toxicity threshold. However, only recently have formalized risk assessment methodologies been developed for the consistent interpretation of toxicological experiments, especially for chronic systemic effects. In the past, the health effects of systemic toxicants have been evaluated on the basis of commonly accepted, but not necessarily well-validated, toxicological practice. Much data interpretation has been thus left to the "expert judgement" of toxicologists, leading to methodological inconsistencies between various regulatory bodies.

The second category of toxic effects arise from *non-deterministic* processes that do not follow the usual dose-response behaviour demonstrated by traditional toxicity mechanisms (Hattis, 1987). Non-deterministic processes do not have an obvious threshold of toxicity, so that exposure to low doses of an agent over prolonged periods of time can give rise to serious toxic effects, which are not necessarily commensurate to the level of exposure. In addition, these processes are irreversible or poorly reversible at the early stages of causation.

This category includes diseases such as cancer and genetic mutation that are the result of *stochastic* processes, where the likelihood of disease occurrence is determined by the probability of a damaging event (hit) to a molecular target in the cell, especially changes in the genetic information contained within the DNA. Stochastic processes act according to a probabilistic pattern of all-or-nothing effects, so that the dose level determines the likelihood of disease occurrence, but not the severity of the disease itself. In principle, stochastic processes do not have a threshold and are therefore not generally considered amenable to the ADI approach.

Other chronic cumulative effects which are not necessarily stochastic may also occur in a non-deterministic fashion through the accumulation over time of many episodic or small-scale damage events such as pulmonary emphysema, noise-induced hearing loss, or neuronal damage due to lead exposure in infants. In addition, many disorders of cell regulation, including most teratogenesis and some carcinogenesis (e.g. tumor promoters), may arise from toxic effects that are influenced more by the timing and pattern of exposure than by dose. However, by taking into account their special mechanisms of action and the unusual exposure kinetics, risk assessments based on the ADI approach for these endpoints may still be attempted, although their validity and accuracy is less certain than for traditional deterministic processes.

2. Acceptable Daily Intake and Safety Factors

A. Development of the Safety Factor Concept

The methodology for determining *acceptable daily intake* (ADI) is based primarily on dose-response experiments in test animals, or less commonly on well-documented exposures in humans. For each type of health effect arising from prolonged exposure to a specific chemical, the ADI methodology examines the dose-response relationship in order to establish a *no-observed-effect level* (NOEL) for a specified biological endpoint (Lu 1983, 1985, 1988; Dourson and Stara, 1983). Due to the uncertainties in interspecies extrapolation of the NOEL from test animals to humans, and because of the individual variability in human sensitivity to the effects of toxic agents, a safety factor must be introduced which reduces the NOEL value to a lower level of exposure that ensures the safety of most individuals in the human population.

Hence the traditional method for calculating the ADI for a toxic agent is to divide the NOEL value by a predefined *safety factor* (SF):

$$ADI = NOEL / SF$$

By convention, the safety factor is most often initially set at 100, which is generally considered sufficient to account for the uncertainties in interspecies extrapolation from animals to humans and for human variations in sensitivity to chemicals (Lu, 1983). However, several other adjustments in the size of the overall safety factor can be introduced according to the nature and quality of the dose-response data if additional toxicological information is available.

The concept of acceptable daily intake (ADI) was first proposed by the Joint FAO/WHO Expert Committee on Food Additives (JECFA) in 1961 (Lu, 1988). The same term was later adopted in 1962 by the Joint FAO/WHO Meeting of Experts on Pesticide Residues (JMPR), and procedures for the determination of ADI values and their application in practice were described. The ADI approach is used by these two international expert bodies for the toxicological evaluation of food additives and pesticide residues, and it has also been directly adopted by the regulatory agencies of many countries in the formulation of national food safety standards (WHO, 1987).

Even before the ADI approach was formalized, the concept of a 100-fold safety factor had been proposed as an "adequate margin of safety" for the derivation of exposure standards for food additives and contaminants by the U.S. Food and Drug Administration (FDA) in 1954. The margin of safety was intended to ensure that substances to be used in food should show no adverse health effect in animals when fed at a dose at least 100 times greater than the likely human dose (Calabrese, 1978a). However, the original safety factor concept was more intuitive than scientific, and it was expressly stated that:

> *...the margin of safety concept is a reasonable approach to the matter, but that its acceptance should not fool researchers and/or the public into believing that there is any experimental or theoretical basis for its existence.*
>
> J.M. Barnes and F.A. Denz (1954), cited in Calabrese (1978a)

The initial rationale for the use of the 100-fold safety factor by FDA was to compensate for intraspecies or interspecies variability to the toxicity of a chemical, to account for sensitive subgroups within the human population, and to adjust for possible synergism between different food additives or contaminants. The selection of the original 100-fold safety factor was based on toxicological studies performed in the late 1940's which suggested a potentially 10-fold greater sensitivity in humans than in test animals, and a 10-fold difference in susceptibility within the human population (Hutt, 1985). Collectively, there would therefore be an expected variation of approximately 100 between test animals and the most sensitive human individuals. The 100-fold safety factor for foods was subsequently adopted by WHO in 1958, and was recommended in 1970 as the conceptual basis for food additive regulation in the United States by the Food Protection Committee of the National Research Council, National Academy of Sciences (Lu, 1983). The FDA later expanded the safety factor concept by recommending the inclusion of an extra 10-fold safety factor in cases where only short term toxicity data were available, resulting in an overall safety factor of 1000-fold for chemicals

with subchronic data for two species, and 2000-fold for subchronic data for one species only (Dourson and Stara, 1983).

In the United States, the ADI approach developed by FDA was subsequently adapted to regulatory standard-setting in other areas of environmental health. Many of the later refinements in the use of safety factors for ADI determination originated from the development of ambient water quality standards for pollutants in drinking water. In 1977, the Safe Drinking Water Committee of the U.S. National Academy of Sciences (NAS) produced a seminal report *Drinking Water and Health* (NAS, 1977), which included significant revisions in the methods for calculating ADIs for national drinking water standards.

The risk assessment recommendations in *Drinking Water and Health* differed from the FDA methods in two respects. First, the NAS proposed that a NOEL should be defined in milligrams per kilogram body weight rather than milligrams per kilogram of diet, since dietary consumption patterns are variable in test animals and differ widely between various human populations. Second, the NAS study explicitly recognized that dose-response data obtained directly from human studies would require only a 10-fold safety factor to account for human variation in sensitivity, whereas animal experimental data would require an overall 100-fold safety factor to account for the additional uncertainties in interspecies extrapolation from animals to humans. Also included was an additional 10-fold safety factor in cases where there was a lack of adequate toxicity data from human and animal studies, producing an overall safety factor of 1000.

In summary, *Drinking Water and Health* made explicit recommendations for the type and magnitude of safety factors that were similar to FDA safety factors (NAS, 1977):

- Valid experimental (epidemiological) results from studies on prolonged ingestion by man with no induction of carcinogenicity.

  (Safety Factor = 10)

- Experimental results of studies of human ingestion not available or scanty (e.g. acute exposure only). Valid results of long-term feeding studies on experimental animals or in the absence of human studies, valid animal studies on one or more species. No indication of carcinogenicity.

  (Safety Factor = 100)

- No long-term or acute human data. Scanty results on experimental animals. No indication of carcinogenicity.

  (Safety Factor = 1000)

Using a methodology derived principally from the 1977 NAS recommendations, in 1980 the EPA published *Guidelines and Methodology Used in the Preparation of Health Effects Assessment Chapters of the Consent Decree Water Criteria Documents* (EPA, 1980). These regulations presented a preliminary set of criteria for the selection of suitable dose-response parameters and safety factors, and a methodology for the determination of the ADI for threshold toxicants in water. The safety factors specified in the 1980 EPA drinking water regulations were essentially the same as those previously delineated in *Drinking Water and Health*.

The ADI methodology was also carried over into other regulatory sectors of EPA, including the derivation of ambient air quality standards and management of toxic wastes. However, the lack of a uniform guideline for ADI usage across the various regulatory offices of EPA lead to serious inconsistencies in the methodological approaches for ADI determination, and in the actual ADI values calculated by different EPA offices for identical chemicals:

> *... program offices have used different approaches for dealing with uncertainty. Some estimate a chronic lifetime RfD (ADI) based on uncertainty factors (safety factors) tied to the available information, and then establish criteria based on that RfD. Some calculate a margin of safety between the highest no-observed-adverse-effect level (NOAEL) of the critical effect and the estimated exposure, and then evaluate that margin specifically in terms of the chemical of interest and its expected exposure pattern. Some offices estimate an appropriate degree of protection on a case-by-case basis, using their best technical and scientific judgement. Finally, some programs have developed their own quantitative techniques for extrapolating or interpolating across data gaps; few of these have yet gained general acceptance within the Agency.*
>
> Preuss and Erlich (1987)

Therefore, throughout the 1980's, EPA attempted to refine the operational definitions for various dose-response measures and safety factors, and to delineate standardized methods for the determination of ADIs, with the intention of imposing methodological consistency throughout EPA regardless of the medium of exposure or regulatory jurisdiction. For example, the usage and validity of safety factors used by EPA for ADI determinations was systematically evaluated by Dourson and Stara (1983). Later, many of the revised safety factors were formally incorporated into the *National Primary Drinking Water Regulations* (EPA, 1985) for determining the ADIs for organic and inorganic chemical contaminants in drinking water sources. Finally, in 1988 the EPA published a formal set of risk assessment guidelines for systemic toxicants entitled *Reference Dose (RfD): Description and Use in Health Risk Assessments* (Barnes and Dourson, 1988). The 1988 EPA Guidelines for systemic toxicants established a consistent methodology for all ADI and safety factor determinations undertaken by EPA for threshold toxicants. They represent an extension of a previously published series of formal EPA Guidelines for the risk assessment of carcinogens, mutagens, and teratogens.

B. NOEL Safety Factor Methodology

1. Definition of ADI

For systemic toxicants that exhibit threshold properties, the acceptable daily intake (ADI) methodology has traditionally been used to derive "safe" levels of exposure. Although the ADI methodology was originally devised for setting safety standards for food additives and pesticide residues, the concept was later applied generically to many different types of standard-setting, including environmental exposures to air or water contaminants and toxic wastes (Dourson and Stara, 1983). The ADI approach is generally applied to exposures over prolonged durations (chronic toxicity), and it attempts to ensure that continuous or frequent exposures will not produce adverse health effects (Lu, 1983). The ADI can be generally defined as:

> *... the daily intake of a chemical (in milligrams per kilogram body weight per day) which, during an entire lifetime, appears to be without appreciable risk on the basis of all known facts at the time.*
>
> Lu (1983)

This definition therefore states a concern primarily about the health effects of chronic lifetime exposure to an agent, and suggests that safety standards will be subject to continual revision as new toxicity data become available. A more explicit operational definition of ADI has been formulated by EPA:

> *An ADI is defined as the amount of toxicant in milligrams per kilogram body weight per day ... which is not anticipated to result in any adverse health effects after chronic exposure to the general population of humans, including sensitive subgroups. Adverse effects are considered as functional impairment or pathological lesions which may affect the performance of the whole organism, or which reduce an organism's ability to respond to an additional challenge*
>
> EPA (1980)

Since no toxicological study based on a restricted number of test animals can ensure absolute safety in humans, the ADI should be expressed in a range from zero exposure to an upper limit of acceptable exposure to a toxic chemical. JECFA has defined the ADI in this manner to emphasize that the ADI represents an upper limit on acceptable risk, and that the lowest technologically feasible exposure levels should be encouraged (WHO, 1987). Similarly, the 1988 EPA Guidelines categorically reject the implication that any exposure less than the ADI is "acceptable" or "safe", or that any exposure in excess of the ADI is always "unacceptable" (Barnes and Dourson, 1988). The ADI concept is viewed as follows:

> *In practice the ADI is viewed by many (including risk managers) as an "acceptable" level of exposure and, by inference, any exposure greater than the ADI is seen as "unacceptable". This strict demarcation between what is "acceptable" and what is "unacceptable" is contrary to the views of most toxicologists, who typically interpret the ADI as a relatively crude estimate of a level of chronic exposure which is not likely to result in adverse effects in humans. The ADI is generally viewed by risk assessors as a "soft" estimate, whose bound of uncertainty can span an order of magnitude. That is, within reasonable limits, while exposures somewhat higher than the ADI are associated with increased probability of adverse effects, that probability is not a certainty. Similarly, while the ADI is seen as a level at which the probability of adverse effects is low, the absence of all risk to all people cannot be assured at this level.*
>
> Barnes and Dourson (1988)

For this reason, the 1988 EPA Guidelines for systemic toxicants have redefined the ADI in terms of the *Reference Dose* (RfD) to avoid any implication of "acceptable exposure" or any "notion of absolute safety":

> *... it should not be categorically concluded that all doses below the RfD are "acceptable" (or will be risk free) and that all doses in excess of the RfD are "unacceptable" (or will result in adverse effects).*
>
> Barnes and Dourson (1988)

The RfD is defined by the 1988 EPA Guidelines as follows:

> *In general, the RfD is an estimate (with uncertainty spanning perhaps an order of magnitude) of a daily exposure to the human population (including sensitive subgroups) that is likely to be without appreciable risk of deleterious effects during a lifetime.*
>
> Barnes and Dourson (1988)

The RfD is calculated in a manner analogous to the traditional ADI methodology, by dividing the no-observed-adverse-effect level (NOAEL) by the overall uncertainty factor (UF) and by an additional modifying factor (MF):

$$RfD = NOAEL / (UF \times MF)$$

Since the methodological basis of the ADI and RfD calculations is essentially similar, the two terms will be considered equivalent in this review, except where otherwise stated.

2. Definition of NOEL and Other Threshold Estimators

The design of toxicological studies normally involves the exposure of groups of test animals to multiple dose levels of a chemical. In practice, every dose-response study should include at least one dose level that exhibits an absence of adverse effects which represents the *no-observed-effect level* (NOEL). In addition, several higher dose levels that exhibit an adverse effect should also be present to define the shape of the

dose-response curve and demonstrate the severity of toxic effects. Normally the NOEL is derived from the most susceptible but appropriate species of animals tested, using the most sensitive indicator of toxicity detected in a long-term (chronic) study (Lu,1988).

In developing guidelines for deriving exposure standards for threshold toxicants, EPA has defined several types of response levels (EPA, 1980; Dourson, 1986; Stara *et al*, 1987). In order of increasing severity, the response levels are defined as follows:

- NOEL
  No-Observed-Effect Level. The highest exposure level producing no statistically or biologically significant increases in the frequency or severity of effects between the exposed population and its appropriate control.
- NOAEL
  No-Observed-Adverse-Effect Level. The highest exposure level producing no statistically or biologically significant increases in the frequency or severity of effects between the exposed population and its appropriate control. *Effects are produced at this level, but are not considered to be adverse.*
- LOEL
  Lowest-Observed-Effect Level; Lowest-Observed-Adverse-Effect Level. The lowest exposure level producing statistically or biologically significant increases in the frequency or severity of effects between the exposed population and its appropriate control. *If the effects produced at this level are not considered to be adverse, the LOEL should be treated as a NOAEL. If the effects produced at this level are considered to be adverse, the LOEL should be treated as a LOAEL.*
- LOAEL
  Lowest-Observed-Adverse-Effect Level. The lowest exposure level producing statistically or biologically significant increases in the frequency or severity of effects between the exposed population and its appropriate control. *The effects produced at this level are considered to be adverse.*
- FEL
  Frank-Effect Level. The exposure level producing unmistakable adverse effects such as irreversible functional impairment or mortality, at a statistically or biologically significant increase in frequency or severity between an exposed population and its appropriate control.

Because the investigator must evaluate whether changes are manifestations of toxicity or merely adaptive responses to exposure, the EPA has adopted specific terms to denote response levels for serious adverse effects: the no-observed-adverse-effect level (NOAEL), and the lowest-observed-adverse-effect level (LOAEL). According to the EPA (1980) definitions, the determination of a NOAEL and LOAEL should consider only genuine toxic effects, and exclude physiological or adaptive effects, such as liver enlargement by stimulation of microsome proliferation or weight loss due to unpalatability of food, which represent a non-pathogenic response to exposure to a test agent:

> *Adverse effects are considered as functional impairment or pathological lesions which may affect the performance of the whole organism, or which reduce an organism's ability to respond to additional challenge.*
>
> EPA (1980)

Under the EPA scheme, a LOEL may actually correspond either to a NOAEL or a LOAEL, requiring a judgement to be made whether the LOEL actually represents a LOAEL or a NOAEL. However, within other regulatory systems, the term NOEL and LOEL can generally be considered to represent significant toxic endpoints exclusively, and can therefore be regarded as synonymous to the NOAEL and LOAEL. For example, for ambiguous toxicity endpoints for food additives, JECFA has advised that *"whenever there is doubt about the significance of a particular effect, it should be considered as an adverse one"* (Lu, 1988).

Based on this dose-response classification system, several guidelines have been formulated for deriving ADI values from animal dose-response data, which are summarized verbatim as follows:

(1) A free-standing FEL is unsuitable for ADI determination.

(2) A free-standing NOEL is unsuitable for derivation of criteria. If multiple NOELs are available without additional data on LOELs, NOAELs, or LOAELs, the highest NOEL should be used to derive an ADI.

(3) A NOAEL, LOEL, or LOAEL can be suitable for ADI derivation. A well-defined NOAEL from a chronic (at least 90-day) study can be used directly, applying the appropriate uncertainty (safety) factor. Caution must be exercised not to substitute FELs for LOAELs.

(4) If, for reasonably closely spaced doses, only a NOEL and a LOEL of equal quality are available, the appropriate uncertainty (safety) factor is applied to the NOEL.

Source: Adapted from EPA (1980).

Since 1980, EPA has made an explicit distinction between the *no-effect level* (NEL) as the dose-response measure defining threshold, and the *no-observed-effect level* (NOEL) as the statistical estimator of threshold. This distinction indicates that the NOEL is not an absolute no-effect level, but rather an experimental approximation of the NEL subject to the influence of sample size and test species sensitivity (EPA, 1980; Nelson, 1984; Crump, 1984). The NEL is therefore located at a dose level lying somewhere between the NOEL and LOEL, although the actual value for the NEL is not precisely known. For widely spaced dose levels, the NOEL may lie at a considerable distance below the actual NEL level, and the LOEL considerably above it.

The NOEL is only an estimate of the actual NEL, because it is derived from a restricted sample of test animals observed under an arbitrary set of experimental conditions. Under different experimental conditions using a more susceptible species or a more sensitive indicator of toxicity, it is conceivable that a lower NOEL value might be obtained.

> *A NOEL is not an inherent property of the animal system but depends upon the experimental design and outcome. Thus it represents, in statistical terms, a statistic or an estimate of the "true no-effect level". This latter term refers to the highest dose which is absolutely safe, and this is an inherent property of the animal system, or in statistical terms, a parameter. For an effect for which no threshold exists, the "true no-effect level" is zero.*
>
> Crump (1984)

Similarly, the estimated NOEL value is determined in part by the statistical sensitivity of the dose-response experiment, which is a function of the sample size of test animals exposed to the agent. Therefore, the existence of a true threshold is not proven by the determination of a NOEL in an animal test system:

> *The existence of a NOEL does not prove or disprove the existence or level of a true threshold; it only defines the highest level of exposure under the conditions of the test that are not associated with a significant increase in effect.*
>
> IRLG (1986)

Therefore, these definitions are limited in their statistical utility because the overall shape and slope of the dose-response curve, and the effect of sample size on statistical confidence intervals, are not normally considered in the determination of the NOEL (Munro and Krewski, 1981; Crump, 1984; Krewski *et al*, 1984). Alternative methods for ADI estimation based on statistical analysis of the entire dose-response curve have been proposed for animal experimental data, but are not yet widely applied in practice (Gaylor, 1983; Crump, 1984; Dourson, 1986).

C. Specification of Safety Factors

1. Definition of Safety Factor

The basic definition for a *safety factor* has been provided by the U.S. National Academy of Sciences:

> *"Safety Factor" or "Uncertainty Factor" is defined as a number that reflects the degree or amount of uncertainty that must be considered when experimental data in animals are extrapolated to man.*
>
> NAS (1977)

More recently, the term uncertainty factor (UF) has been introduced by EPA as a substitute for safety factor (SF), in order to avoid implications of absolute safety at the ADI levels calculated by applying safety factors to NOELs obtained from animal experiments (EPA, 1980, 1988; Dourson and Stara, 1983):

> *The term "safety factor" suggests, perhaps inadvertently, the notion of absolute safety, i.e. absence of risk. While there is a conceptual basis for believing in the existence of a threshold and "absolute safety" associated with certain chemicals, in the majority of cases a firm experimental basis for this notion does not exist.*
>
> Barnes and Dourson (1988)

For consistency, the term safety factor (SF) will be used throughout this review as synonymous with uncertainty factor (UF).

Table 2.1
Guidelines and Experimental Support for the Use of Uncertaintly (Safety) Factors*

(adapted from Dourson and Stara, 1983)

Guidelines	Experimental Support
Use a 10-fold factor when extrapolating from valid experimental results from studies on prolonged ingestion by man. This 10-fold factor protects the sensitive members of the human population estimated from data garnered on average healthy individuals.	Log-probit analysis; Composite human sensitivity
Use a 100-fold factor when extrapolating from valid results of long-term feeding studies on experimental animals with results of studies of human ingestion not available or scanty (e.g., acute exposure only). This represents an additional 10-fold uncertainty factor in extrapolating data from the average animal to the average man.	Body-surface area dose equivalence; Toxicity comparison between species
Use a 1000-fold factor when extrapolating from less than chronic results on experimental animals with no useful long-term or acute human data. This represents an additional 10-fold uncertainty factor in extrapolating from less than chronic to chronic exposures.	Subchronic/ chronic NOEL or LOEL comparison
Use an additional uncertainty factor of between 1 and 10 depending on the sensitivity of the adverse effect when deriving an ADI from a LOEL. This uncertainty factor drops the LOEL into the range of a NOEL.	LOEL/NOEL comparison

*These factors are to be applied to the highest valid NOAEL or NOEL which does not have a valid LOEL equal to or below it, in calculating an ADI when no indication of carcinogenicity of a chemical exists.

2. Selection of Safety Factors

According to the 1988 EPA Guidelines (Barnes and Dourson, 1988), four standard safety (uncertainty) factors are defined collectively (Table 2.1), each with a default value of 10:

- **human variability** (UF = 10H)
- **interspecies extrapolation** (UF = 10A)
- **adjustment for less-than-chronic studies** (UF = 10S)
- **adjustment for the use of LOEL** (UF = 10L)

Under this scheme, the traditional 100-fold safety factor has been formally disaggregated into its two principal components. The first standard 10-fold safety factor is introduced to account for *human variation* in sensitivity to chemicals (10H). The second standard 10-fold safety factor accounts for the uncertainties in *interspecies extrapolation* from test animals to humans (10A). To adjust for the use of a NOEL obtained from a *subchronic* (90 day) animal study, in the absence of comparable data from a chronic (180-360 day) study, an additional standard 10-fold safety factor is included for subchronic studies (10S). Finally, if the NOEL cannot be determined directly from dose-response data because it is located below the observed dose range of the animal experiments, then the *lowest-observed-effect level* (LOEL) is substituted for the NOEL by including an additional 10-fold safety factor to adjust the LOEL into the range of the NOEL (10L).

In simple cases, the overall safety factor can be calculated by multiplying the standard safety factors, selected according to the criteria listed in Table 2.1. Collectively, the standard safety factors produce an order-of-magnitude value for the overall safety factor, such as 10, 100 or 1000:

$$SF_{TOTAL} = \text{product } (SF_i) = SF_1 \times SF_2 \times SF_3 \times SF_4$$

The 1988 EPA Guidelines also provides for the inclusion of a *Modifying Factor* (MF) which serves as a subjective adjustment factor to account for any additional scientific information that may affect the value of the ADI, based on the "best professional judgement" of the risk assessors. The rationale for the inclusion of a modifying factor is articulated as follows:

> *Safety factors do not pretend to have mathematical precision. When you have very little information, order-of-magnitude safety factors may be about the best you can do in assessing the risk from low levels of exposure. The use of an elaborate mathematical model in such situations may convey a misleading impression of precision. On the other hand, in situations where you do have more information, where there is, for example, a known mechanism for bioaccumulation or enzyme measurements showing how much of the enzyme is being metabolized, then I think it is vital that such information be taken into account in the assessment process. In such situations it would be wrong to ignore the data and to use a simple procedure that mechanically selects a certain power of ten as the appropriate safety factor.*
>
> EPA Science Advisory Board (EPA, 1985)

The MF value may be any number greater than zero up to a maximum of 10, with a default MF value of one. This allows for the calculation of an adjusted overall safety factor that is not always a power of 10. For example, a dose-response study with a large sample size (250 animals) could produce a subjective adjustment factor (MF) of 0.8, which if multiplied by a standard UF of 1000 would result in an adjusted overall safety factor of 800 (Barnes and Dourson, 1988).

The 1988 EPA Guidelines state that the principle adjustments covered by the MF concern the completeness of the overall database and the number of species tested, but they do not otherwise delineate decision criteria or specify the magnitude of adjustment values. Some additional criteria for the application of a modifying factor have previously been described in the *National Primary Drinking Water Regulations* (EPA, 1985):

- less-than-lifetime exposure of test animals
- less-than-lifetime study (presumably for pathological follow-up)
- study based on different route of exposure (e.g. food rather than water)
- chemical is an essential nutrient at low levels, toxic at higher levels
- endpoint was much more sensitive than traditional endpoints used for ADI calculations.

The first three of these criteria would give rise to an MF greater than one, thereby increasing the size of the adjusted safety factor, and decreasing the ADI value accordingly. The last two criteria would result in an MF less than one, producing a smaller adjusted safety factor and larger ADI than otherwise would be the case.

Based on such criteria, *adjusted safety factors* with values of less than 10 are possible. For example, in the case of chemicals with well-characterized responses in sensitive individuals, such as supplemental fluoride in drinking water, an MF of 0.1 could be applied to an existing safety factor of 10 to produce an adjusted safety factor as small as one (Barnes and Dourson, 1988). However, this example may be misleading, since the selected endpoint for fluoride is tooth mottling, which is a virtually harmless cosmetic effect. The minimum dose of fluoride required to produce serious health effects lies in a considerably higher range.

In cases where the data did not precisely fulfill the conditions for one category of safety factors, earlier EPA practice also allowed the size of the overall safety factor to be modified by the use of *intermediate safety factors*, according to non-standard criteria such as the severity of the toxic endpoint and the dose-response characteristics of the agent (EPA, 1980). Such an intermediate safety factor could be

located along a logarithmic scale (e.g. 32 is half-way between 10 and 100 on a log scale). In addition, the LOEL safety factor was permitted to vary over a range of values between one and 10 based on judgmental considerations, such as the steepness of the dose-response slope (EPA, 1985). Agents with a steep dose-response slope are more likely to have a LOEL in closer proximity to the NOEL than are agents with a shallow slope. When the ADI was derived using a LOEL, the overall safety factor would therefore lie between 100 and 1000, and a smaller safety factor correction would be required for steep dose-response slopes than for shallow slopes.

However, the more recent 1988 EPA Guidelines have incorporated such subjective adjustments into the Modifying Factor, rather than attempting to adjust the value of the individual safety factors as previously described. Nevertheless, the 1988 EPA Guidelines have conspicuously avoided any detailed decision criteria for MF adjustments based on subjective scientific considerations. Possible candidates for such a criteria list would include: analysis of the statistical variability of the NOEL and the underlying UFs, characterization of the dose-response curve (including slope), and the severity of biological endpoints. Formal inclusion of these criteria into the calculation of the MF and the adjusted safety factor are currently under active development by EPA scientists.

D. Validation of Safety Factors

1. Overall Safety Factor

A review by Bigwood (1973), working on behalf of WHO, offered an expanded rationale for the 100-fold safety factor, by analyzing the different sources of variation between species, and assigning a certain portion of the 100-fold safety factor to them. The five sources of variation were identified as:

- body size (test animals versus humans)
- food consumption patterns
- water balance (consumption)
- hormonal functions and their effect on food intake
- differential inherent sensitivity to toxic effects.

When approximate values based on existing toxicological data were arbitrarily assigned to each of the first four categories, it was estimated that a safety factor of about 60-fold was required to account for the first four factors. Taking into consideration the fifth factor, which was largely unquantifiable, it was then proposed that all five factors could be reasonably accounted for by extending the safety factor from 60 to 100. It was therefore concluded that a safety factor of 100 possessed a high degree of validity, with excursions above or below this value (e.g. 50 to 500) being unnecessary in most cases. Bigwood concluded that adjustments to the 100-fold safety factor should be made only in unusual circumstances.

2. Specific Safety Factors

To provide theoretical and empirical support for the adequacy of each of the 10-fold safety factors prescribed in the 1988 EPA Guidelines, Dourson and Stara (1983) have undertaken a systematic validation study based on pre-existing sets of animal toxicological data.

a) Intraspecies Variation

The 10-fold safety factor for variation in intraspecies sensitivity was validated based on an analysis of dose-response data from 490 separate probit vs log-dose slopes obtained from acute lethality (LD_{50}) experiments. A sensitive subpopulation was defined by scaling down the median response (LD_{50}) by a 3-probit reduction, corresponding to a cumulative response of 0.13% ($LD_{0.13}$). The log-dose: probit slope of the extrapolation curve connecting the LD_{50} and $LD_{0.13}$ points was then used as a measure of the intraspecies variation in sensitivity between the average (median) and most sensitive members of a population.

Steep curves with slopes greater than three indicate that a 10-fold safety factor would reduce the applied dose below the $LD_{0.13}$ level, and therefore would adequately protect sensitive subpopulations. Shallow curves with slopes of less than three indicate that a 10-fold safety factor would be insufficient to reduce the applied dose below the $LD_{0.13}$ level, and consequently would not protect the most sensitive members of the population. It was found that 92% of the probit vs log-dose slopes had values greater than three, which demonstrates that for most chemicals a 10-fold safety factor is sufficient to account for intraspecies variation in sensitivity.

While the results of this study are reassuring, any intraspecies validation study using animal experimental data will have only indirect relevance to humans. The low variation in responsiveness observed in lab animals is a reflection of their genetic homogeneity and controlled living conditions. Genetically heterogeneous humans, who also live under more varied environmental conditions, will probably exhibit a broader variation in sensitivity to chemicals than lab animals.

Calabrese (1985) reported large interindividual variations in human pharmacokinetic and DNA damage/repair responses to environmental chemicals. Although the vast majority of people studied responded within a factor of 10, human variation in the most sensitive individuals often exceeded a factor of 10, and occasionally approached a factor of 1000. Depending on the chemical, the 10-fold safety factor for human variability appears to provide adequate protection for approximately 80-95% of the population, while the remaining 5-20% would respond by a factor of greater than 10. Moreover, since the variation caused by each pharmacokinetic factor was assessed individually, the combined effects of several pharmacokinetic polymorphisms acting in concert within hypersensitive individuals may be considerably greater than otherwise anticipated from single factor studies (Calabrese, 1978b).

Hattis *et al.* (1987) have reported similar findings in a biostatistical analysis of human variability, using measurements of interindividual variability for several pharmacokinetic parameters, including elimination half-lives, maximal blood concentrations, and AUC (area under the curve). From datasets

for 101 chemicals, a secondary analysis of the statistical (log-normal) distribution of these measures revealed that the 10-fold safety factor was an appropriate default value, but in special cases this value might be inadequate. These findings also take into account only one aspect of human interindividual variation in sensitivity, pharmacokinetic variability. Other factors such as variable exposure levels and differing toxicodynamic responsiveness to toxic agents have not been examined in these studies. A source of concern is that much of the safety factor for human variation is taken up in accounting for differences solely in pharmacokinetic patterns.

b) Interspecies Extrapolation

Since test animals are much smaller than humans, the transformation of an ADI established by animal tests to humans requires a conversion factor for body size (Dourson and Stara, 1983). Many possible scaling factors can serve as the basis of this transformation procedure, but body surface area is often considered as a more valid measure for interspecies conversion than body weight, because it accounts better for interspecies variation in metabolic rate (Calabrese, 1983, 1987; Vocci and Farber, 1988). It is assumed that different species are equally sensitive on the basis of dose per unit surface area, because the relative rates of detoxication or elimination of chemicals are accounted for in this procedure. An indirect calculation of surface area is usually made according to the 2/3 power of body weight. The generic interspecies adjustment factor for scaling is then calculated as the cube root of the ratio of human body weight (70 kg) to animal weight W:

$$\text{interspecies adjustment factor} = \sqrt[3]{70/W}$$

This conversion factor is supposed to produce an *equitoxic dose* in the test animals and humans. The human dose is derived by dividing the animal dose (mg/kg body weight) by the interspecies adjustment factor. By applying the scaling formula to data obtained from previous interspecies sensitivity studies, the interspecies adjustment factor has been found to range between six and 12, which is in accord with interspecies differences in susceptibility reported by the NAS study (NAS, 1977).

Several empirical scaling studies are cited as supporting the notion that the interspecies adjustment factor usually falls within the 10-fold range that is used as the safety factor for interspecies extrapolation (Dourson and Stara, 1983). However, many of these studies are of questionable relevance and validity, for example, studies comparing the interspecies sensitivity of animals and humans for a series of antitumor alkylating drugs. It is also disturbing to realize how large a proportion of the 10-fold interspecies extrapolation safety factor is taken up in accounting merely for body size differences. For example, the rat-to-human adjustment factor is estimated by Dourson and Stara as approximately six. There is little residual margin of safety remaining for other interspecies factors, such as differences in pharmacokinetic rates and intrinsic species sensitivities that are not a function of body size.

Moreover, there is serious disagreement with the surface area concept among different regulatory bodies. The body surface adjustment has been considered but rejected by JECFA and JMPR, on the grounds that metabolism of xenobiotics is not necessarily well correlated with general body metabolic rates, and that the surface area conversion is not well validated by experimental data (Lu 1985, 1988). For a series of 28 organophosphate pesticides whose dose-response data are well characterized, it has been demonstrated that surface area conversion does not account adequately for differences in interspecies sensitivity between the dog and rat. Lu concludes that body surface conversion should not be considered better than body weight as a basis for interspecies extrapolation.

Until recently, the surface area species conversion factor has been used by EPA only for estimating lifetime risk for exposures to carcinogens. Now EPA is considering using this approach for both carcinogens and systemic toxicants, although no official guideline has yet been published (Stara *et al.*, 1987). ADI procedures in other countries (e.g. Dutch Health Council) have allowed a reduction in the 100-fold safety factor to a 30-fold safety factor for systemic toxicants when surface area conversion is included in the ADI calculation (Vocci and Farber, 1988). However, given the persistent disagreement over the surface area conversion, such modification in the size of the safety factor is questionable.

c) Subchronic to Chronic Adjustment

To determine the effect of using data from subchronic dose-response studies to determine NOEL values, Dourson and Stara (1983) compared NOELs obtained from chronic (2-year) dose-response studies to NOELs obtained from subchronic studies that ranged between 30 to 310 days (average 92 days). LOEL values were also compared in a similar fashion. The NOEL or LOEL values for various chemicals were obtained from a series of toxicity studies performed by Weil and McCollister (1963). The ratio of the subchronic NOEL (or LOEL) to the chronic NOEL (or LOEL) indicates the degree of overestimation of the NOEL (or LOEL) when subchronic studies are used, and the corresponding overestimation of the ADI value. Based on comparison of 30 studies with NOEL data and 22 studies with LOEL data, it was found that the NOEL or LOEL values obtained by subchronic and chronic studies differed by less than a factor of two in over half of the cases. In 96% of the studies, the subchronic/chronic NOEL ratios were less than 10. These findings support the use of a 10-fold safety factor to account for the uncertainties of NOEL estimation based on subchronic studies, with a relatively high degree of confidence. However, the possibility of missing latent or subtle pathological changes in subchronic tests cannot be entirely excluded.

d) LOEL to NOEL Adjustment

Dourson and Stara (1983) also examined the effects of using the LOEL instead of the NOEL to calculate the ADI, using the same toxicity studies as previously described (Weil and McCollister, 1963). For each subchronic or chronic study, the ratio of the LOEL to the NOEL can be viewed as the reduction in the LOEL that would be required to produce a corresponding NOEL. The ratio thereby measures the degree

of overestimation of the ADI that occurs when LOELs are substituted for NOELs. It was found that virtually all chemicals examined had LOEL to NOEL ratios of less than 10. For 96% of the chemicals the LOEL to NOEL ratio was less than 5. These findings support the view that a 10-fold safety factor is sufficient to account for the use of LOEL values, when NOEL values are not available for a chemical. However, this assertion holds true only for well designed multiple-dose experiments, where there is an adequate assurance that the LOEL lies in approximately the same dose range as the NOEL.

3. Alternative Approaches to Safety Factor Methods

A. Criticisms of the NOEL-SF Method

Compared to the sophisticated statistical techniques that have been devised for quantitative risk assessment of carcinogenic agents, the characterization of the dose-response relationship in conventional NOEL-SF methods is severely limited. Consequently, the NOEL-SF approach has been criticized for several methodological weaknesses, which are considered by many statisticians to compromise the accuracy and usefulness of the ADI estimates (Munro and Krewski, 1981; Crump, 1984). These weaknesses include the following:

- misuse of the NOEL data point

The NOEL value is generally considered useful for ADI determination only when it represents one of a series of data points in a well-characterized dose-response curve (EPA, 1980). A free-standing NOEL data point cannot be considered useful because there is no indication of its proximity to the LOEL. Consequently a free-standing NOEL may be located several orders of magnitude below the threshold region. The NOEL may be also considerably underestimated if data points are few and widely separated. Therefore an adequate number of closely spaced dose points (generally a minimum of 3-5 points) are considered necessary for more precise determination of the ADI as estimated from the NOEL.

- effects of sample size

The statistical sensitivity of a dose-response experiment will be determined mainly by the number of animals used in the study. Studies with large sample sizes have a better chance of detecting a statistically significant result than smaller studies. Response rates of 0/10 and 0/1000 obviously have different statistical interpretations, yet in principle both can be used to derive a NOEL (Munro and Krewski, 1981). Unlike quantitative risk assessment for carcinogens, where 95% upper confidence limits are routinely calculated and used as a conservative upper bound estimate of risk, the NOEL is a statistic that represents the best estimate of risk without any explicit reference to upper-bound risk (Crump, 1984). Hence the NOELs and the derived ADI values obtained in large-scale experiments will tend to be smaller and more restrictive, while studies with small numbers of animals will have large and more generous ADIs. Thus there is a direct disincentive to undertake studies with large numbers of animals, because these are more likely to produce lower NOEL estimates and more stringent ADI exposure limits. Minimum sample sizes must therefore be specified in the design of dose-response experiments to ensure that NOELs and ADIs are not seriously overestimated.

- characterization of dose-response

In the conventional NOEL-SF approach, the NOEL (or LOEL in the absence of a NOEL) is the sole data point in the dose-response curve that is utilized for the derivation of the ADI (Crump, 1984). Apart from ensuring that the number and spacing of data points is adequate to provide a reasonable estimate of the NOEL, all other data points in the curve are ignored. Hence the NOEL-SF estimation procedure usually ignores the slope and shape of dose-response curve (Munro and Krewski, 1981, Crump, 1984). Therefore valuable information about the overall downward trend of the dose-response curve is lost, and it is not possible to provide a more accurate estimate of the "true no-effect level" (NEL), or to produce a low-dose extrapolation curve for estimating the *benchmark dose* that corresponds to a predefined level of risk.

In cases where a steep slope is found, the dose-response curve extrapolates to the zero-response level quite rapidly, and therefore a relatively high NEL value can be predicted from the slope. Conversely, a shallow slope extrapolates to zero response more slowly, so that a lower apparent NEL value is obtained. Because shallow slopes have a wider distance between the NOEL and LOEL points than steep slopes, the threshold dose could be considerably underestimated for shallow slopes when only the NOEL point itself is determined.

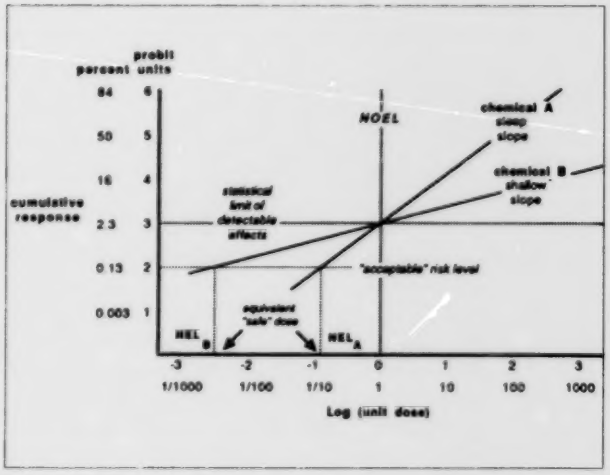
For example, if two chemicals (*A* and *B*) have apparently identical NOELs but different dose-response slopes, the chemicals will exhibit a large difference between their respective benchmark doses within the low-dose region of the curve (Figure 3.1). Assuming that a response level of 2.3% (equivalent to 3 probit units on a probit:log-dose plot) falls just below the detectable limit of statistical significance for the two chemicals, the NOEL is derived for both chemicals as one dose unit from the ×-coordinate of the dose-response curve at the 2.3% (3 probit) response level. However, if the "acceptable"

risk level is defined as a 0.13% response (2 probits), representing a no-effect level for 99.87% of the population, then the equivalent "safe" dose (benchmark dose) would be approximately 1/10th of a dose unit for chemical *A* (steep slope), but 1/300th of a dose unit for chemical *B* (shallow slope). Their respective NELs (which do not show on a probit plot) would also be widely divergent.

- arbitrary values for the safety factor

The present values assigned to the various safety factors have originated from a combination of historical accident and rational assessment of the sources of uncertainty in deriving ADIs. *"The fact that humans have 10 fingers undoubtedly played a role in the specific factor selected"* (Crump, 1984). The 1988 EPA Guidelines that delineate the magnitude and usage of the standard safety factors do not provide an explicit rationale or methodology for adjusting these safety factors according to sample size, number and spacing of data points, or slope of the dose-response curve, although these factors are briefly listed under criteria for establishing the modifying factor (MF). In addition, the NOEL-SF method provides no procedure for predicting levels of risk for doses below the NOEL, or for calculating benchmark doses that correspond to predefined levels of risk. In order to redress some of these criticisms, which essentially relate to the statistical characterization of the dose-response curve, a number of new safety factor models have been recently proposed, based largely on analogy to mathematical curve-fitting models previously developed for quantitative risk assessment of carcinogens.

Figure 3.1
Effect of Slope on Risk at Low Dose



B. Mathematical Safety Factor Models

1. Normal Distribution (Probit) Models

The earliest attempts to account for the slope of the dose-response curve in ADI determinations were based on the statistical properties of the log-normal distribution of dose-response encountered in most animal experiments. As early as 1956, it was proposed that a safety factor could be derived by measuring the slope of the log dose-response curve and its standard deviation. The ADI could be derived by determining the threshold dose (i.e. NOEL) and extrapolating downward by six standard deviations (Vettorazzi, 1976). In effect, the ADI was calculated as the NOEL minus 6 probit units, where the subtraction of probits represented the safety factor. In 1961, a more sophisticated extrapolation model based on the log-normal distribution was proposed by Mantel and Bryan for carcinogens, although the model was equally applicable (indeed more suitable) for threshold toxicants. The Mantel-Bryan model proposed that a low dose extrapolation could be performed to calculate the dose producing a benchmark risk of one in 100 million ("virtual safety"), using a log-probit model with an arbitrarily shallow probit slope of one standard deviation per unit log-dose (Mantel, 1975). Although this model was adopted for non-threshold carcinogens in the early 1970's by some regulatory bodies, it was not generally accepted for the calculation of ADIs for threshold systemic toxicants.

The NAS study *Drinking Water and Health* (NAS, 1977) also reviewed the used of log-probit and log-logistic models for the determination of exposure limits for contaminants acting as threshold toxicants in drinking water. Although the models could be used to estimate risk for a toxic effect, they required data from lifetime feeding studies with large numbers of animals, in order to properly define the normal distribution response curve for the toxicant. Since sufficient data was rarely available for this purpose, the NAS continued to recommend the NOEL-SF approach for ADI determination except in cases where adequate dose-response data was available (Cotruvo, 1988).

2. Gaylor Linear Extrapolation Method

One of the first mathematical safety factor methodologies was proposed by Gaylor (1983). The underlying premise of the Gaylor procedure was that the safety factor represents in part an uncertainty factor that is a function of sample size in the dose-response experiment. The traditional safety factor of 100 may be either too large or too small, depending on the sample size employed. A small sample size will produce a large upper confidence limit (UCL) for the estimated NOEL, while a large sample size will produce a small UCL for the same NOEL. Hence the size of the safety (uncertainty) factor should be

proportional to the UCL of the NOEL, and inversely proportional to the desired level of tolerable risk. Risk can be predefined as 10^{-4} or any other specified level of risk. The Gaylor method proposed that the safety factor could be calculated as:

$$F = U/R$$

where *F* is the safety factor, *U* is 95% UCL of response at given dose level, and *R* is the specified risk level.

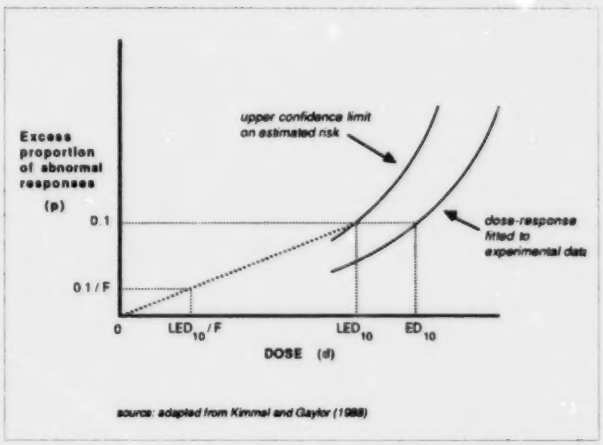
In many cases, the NOEL lies far along the lower end of the dose-response curve and cannot be accurately estimated by direct observation. The Gaylor procedure instead selects the ED_{01}, the dose producing a 1% toxic response rate, as an artificial *reference dose level* or *"pseudo-NOEL"*. Although the ED_{01} is not often directly observable because it is too close to the background response rate, the upper confidence limit of the ED_{01} can be readily calculated from the dose-response data using a mathematical curve-fitting model. Any conventional low-dose extrapolation model could be used, with the appropriate model selected on the basis of goodness-of-fit to the observed slope, or on the basis of plausible biological mechanism. Since the shape of the dose-response curve is often upwardly concave at low dose, the UCL of the ED_{01} generally provides a more accurate estimate for the value of F than the UCL of the NOEL.

The value for *U* can be calculated from the fitted dose-response curve as the 95% UCL of the estimated risk at ED_{01}, which is the equivalent of a hypothetical LOEL. The value for *U* is then divided by the desired risk level *R* (e.g. 10^{-4}) to calculate the safety factor *F*. A straight line can then be drawn from the point on the upper confidence limit curve corresponding to 1% response level to the origin of the graph. Based on this linear extrapolation of the dose-response curve, an ADI can be calculated by dividing the ED_{01} dose by the value for *F*:

$$ADI = ED_{01}/F$$

However, a more recent version of the Gaylor method recommends the use of the ED_{10}, the effective dose corresponding to an excess risk of 10%, to serve as the reference dose level in place of the ED_{01} (Kimmel and Gaylor, 1988). The ED_{10} is the preferred reference dose because its estimation does not require extrapolation of the fitted curve much below the normal range of observable dose-response data. The curve corresponding to the upper confidence limit on estimated risk is also obtained in order to determine the LED_{10}, which is defined as the lower confidence limit on dose for the ED_{10}. The LED_{10} therefore represents the dose where a benchmark risk of 0.1 intersects the curve for the upper confidence limit on risk (Figure 3.2). A benchmark dose for any predefined level of risk can subsequently be calculated by selecting a safety factor *F* for the LED_{10} and dividing the LED_{10} by *F*. For example, to achieve a risk level of 10^{-4}, a safety factor of 1000 would be applied to the LED_{10} (i.e. $0.1/1000 = 10^{-4}$). Since risk is a linear function of dose in the region below the LED_{10}, the benchmark dose for any specified small level of risk can be derived from the calculation of the LED_{10}/F, and used as an ADI. This procedure is very similar to the linear interpolation technique previously devised by Gaylor for analysis of carcinogenesis dose-response data (Gaylor and Kodell, 1980).

Figure 3.2
The Linear Extrapolation Method Using the Upper Confidence Limit on the ED_{10}



Using yet another approach, Gaylor (1988) has advocated adopting the multistage model of carcinogenesis as the basis for a low-dose extrapolation model for systemic toxicants, in a manner analogous to its use for non-threshold carcinogens. The principal advantage of this extrapolation model is that its general sigmoidal shape makes it useful for describing the dose-response of many non-carcinogens, which generally exhibit an upwardly concave dose-response curve in the low-dose region. The multistage model can be defined by a polynomial equation describing the probability of response $P(d)$ as a function of the dose d, which at low doses becomes a simple linear dose-response function. Using the fitted curve, a reference dose can then be estimated from the lower confidence limit on the dose expected to produce 1% or 10% disease incidence (the LED_{01} or LED_{10}) and a benchmark dose for a predefined level of risk can be calculated by applying a safety factor to the reference dose as previously described. When available, pharmacokinetic information may also be incorporated into the model to predict the effective dose of the toxicant at the target organ, and thereby reduce the magnitude of the estimated risk.

The principal advantage of the Gaylor approach is that it produces a less subjective estimate of low-dose risk than NOEL-SF methods. Its specific advantages are as follows: (1) it allows for the flexible determination of the safety factor, (2) it takes into account the shape and slope of the dose-response curve, (3) it provides a conservative low-dose extrapolation and therefore produces quantitative values similar to the safety factor approach, (4) it requires no assumptions about the actual shape of the dose-response curve in the low-dose region, and (5) the relative magnitude of risk calculated from the NOEL-SF procedure compared to the

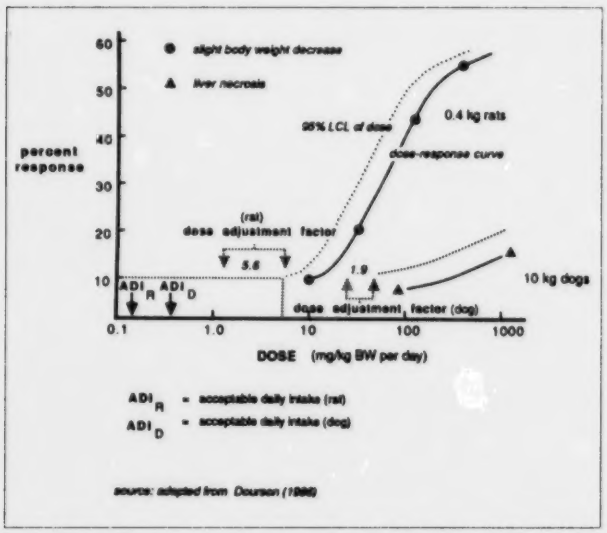
mathematical extrapolation procedure serves as an independent validation of the ADI derived from NOEL-SF calculations.

The main disadvantage of the Gaylor method is that it is applicable only to quantal (binary) data, and it requires more than the usual number of data points to adequately define the shape of the dose-response curve for the mathematical curve-fitting procedure. Also, it is only relevant to the experimental animal model, so an additional safety factor must also be introduced to account for interspecies extrapolation to humans. Another problem is how to determine what is an acceptable level of risk (e.g. 10^{-4}, 10^{-6}, etc.) for deriving benchmark dose, and how to relate this value to the concept of threshold.

3. Dourson Safety Factor Method

Another procedure for deriving flexible safety factors has been developed by Dourson (1986). This method has a similar conceptual approach to that of Gaylor, but the details differ in several ways. The Dourson procedure also may select the ED_{10} as a reference dose level, although any reasonable reference dose level is also acceptable. For the calculation of the ADI, the dose-response curve is fitted using the mathematical model of choice, and the 95% LCL of the dose associated with the ED_{10} risk level is determined in a manner similar to Gaylor (Figure 3.3).

Figure 3.3
The Dourson Safety Factor Method for ADI Determination



The ADI is then calculated by dividing the LCL of the ED_{10} (i.e. LED_{10}) by several safety factors. First, a dose adjustment factor is included for interspecies extrapolation, normally based on a cube-root surface area conversion factor. Then a safety factor of 10 is used to account for human variation, and a second safety factor of one to 10 is used as an adjustment for using a LOEL value. The latter factor is included because Dourson considers the ED_{10} to be the equivalent of an experimental LOEL. The precise value of the second safety factor is determined by the slope of the dose-response curve and the severity of the toxic effect. For a steep slope with a mild toxic effect the SF would equal one; for a shallow slope with a severe toxic effect the SF would equal 10. The overall safety factor will therefore range between 10 and 100.

The Dourson method is basically an extension of the traditional NOEL-SF approach developed by EPA, with the single-point NOEL being replaced by a reference dose, represented by the 95% LCL of the ED_{10} derived from a curve-fitting model. The principal advantage of this approach is that it avoids dependence on the single-point NOEL, and instead employs all of the experimental data to define the shape and error bounds of the dose-response curve. Factors such as slope and severity of effect are also taken into account. However, in contrast to Gaylor, the Dourson method adheres to a conventional safety factor approach for ADI calculation, and does not introduce a linear extrapolation for estimating risk at low doses. Thus only a single ADI can be estimated, and no calculation of benchmark doses for predefined risk levels is possible.

4. Crump Benchmark Dose Method

Crump (1984) has proposed a method for the derivation of the ADI which is based on analogy to existing dose-response models for carcinogens. The operational basis for his model is similar to that of Gaylor (1988). In place of the conventional NOEL-SF, the ADI is calculated in terms of a benchmark dose (BD), defined as *"the statistical lower confidence limit on the dose corresponding to a small increase in effect over the background level"* (Crump, 1984). The "just-above backgound" response point used to define the BD would typically be located near the lower end of the observable dose-response curve, with an increase in response small enough to reflect the shape of the dose-response curve, but large enough to ensure that the lower confidence limit will not be too dependent on the mathematical model selected for its calculation. A response level of 1% to 10% above background would satisfy these requirements. Typically a 10% response level would be selected, so that the BD is functionally equivalent to the LED_{10} of the Gaylor (1988) method. The Crump method provides for the application of several different curve-fitting models. Provision is made both for quantal data, where only the presence or absence of a response is noted, and for continuous data for which the severity of a biological effect is measured.

For *quantal* dose-response data, Crump proposes the use of a *quantal linear regression* (QLR) model where the probability of effect $P(d)$ varies as a function of dose d, according to the expression:

$$P(d) = c + (1-c)[\,1 - \exp[-q_1(d-d_o)]\,]$$

where c is the constant for background response rate for a critical endpoint, d_o is the threshold dose, and q_1 is the slope parameter for the linear dose-response function. Ignoring background, this equation reduces to:

$$P(d) = 1 - \exp[-q_1(d-d_o)]$$

which is a slightly modified form of the linear (1-hit) model used for low-dose extrapolation of carcinogenesis data, except that the dose variable is redefined as the incremental dose above a threshold dose. Therefore, when $d < d_o$, this defines the range of doses below the threshold where the response rate above background will be zero.

Similarly a *quantal polynomial regression* (QPR) model can be defined (ignoring background) by an exponential polynomial equation of the form:

$$P(d) = 1 - \exp[-q_1(d-d_o) - q_2(d-d_o)^2 ... - q_k(d-d_o)^k]$$

which is an expression normally used to describe the multistage model of carcinogenesis, modified by the introduction of a threshold-dose parameter d_o. Such a polynomial expression allows for the fitting of curvilinear dose-response data, and is therefore more flexible than the linear model. Crump also provides expressions for models describing the log-normal distribution and the Weibull distribution.

For *continuous* dose-response data, Crump suggests curve-fitting models that are analogous to those for quantal data. He proposes a *continuous linear regression* (CLR) model of the form:

$$m(d) = c + q_1(d-d_o)$$

where $m(d)$ is the degree of effect above a baseline value c, expressed as a simple linear function of the dose that exceeds the threshold dose d_o. A *similar continuous polynomial regression* (CPR) model can be expressed as:

$$m(d) = c + q_1(d-d_o) + q_2(d-d_o)^2 ... + q_k(d-d_o)^k$$

Once the dose-response curve for a selected regression model and its statistical confidence limits have been generated, the benchmark dose (BD) can be readily calculated. The BD is used in place of the conventional NOEL, and is defined as the lower confidence limit on a dose that corresponds to a specified increase above background in the level of extra risk for quantal data or the amount of extra effect for continuous data.

For *quantal data*, the BD is defined as the dose d that produces a specified level of extra risk, defined as:

$$[P(d) - P(0)] / [1 - P(0)]$$

Extra risk can be interpreted as the probability of the occurrence of an effect at dose d, above the background response rate $P(0)$ in the absence of dose.

For *continuous* data, the BD is defined as the dose d that produces a specified level of *extra response*, defined as:

$$[m(d) - m(0)] / m(0)$$

which is the amount of change in the magnitude of an effect at dose d relative to the baseline level of the effect.

Once the BD has been determined, a safety factor must be introduced to bring the predefined benchmark risk from the 1% or 10% level downward to a level of "acceptable" risk. For example, if the BD has been set at a predefined risk level of 1%, then a 100-fold safety factor would be required to reduce the BD to a risk level of 10^{-4}. However, Crump does not directly specify the nature or magnitude of safety factors for the BD. Rather he suggests that safety factors could be determined empirically, by comparing the values for ADIs obtained by the NOEL-SF method with BD values for the same chemicals. Given a sufficient number of comparisons, the appropriate safety factor for the BD could be determined as the average of the BD/ADI ratios. Thereafter this safety factor could be routinely applied to all BDs, except in specific cases where special safety factors were required.

Alternatively, the curve-fitting models could be used to extrapolate the dose-response curve to a low-dose region to provide a direct estimate of the threshold dose, and benchmark doses corresponding to specified levels of risk could be estimated using the extrapolated curve. For regulatory standard-setting purposes, the latter approach would be especially useful in calculating cost-benefit ratios for modeling alternative exposure scenarios, which cannot be done using the conventional NOEL-SF approach. However, the uncertainties of the shape of the dose-response curve at low dose and the sensitivity of the extrapolated curve to model selection basically discourage this approach. Threshold estimates will tend to differ widely depending on the model selected, in a manner analogous to differences in the estimated values for *virtually safe dose* (VSD) for carcinogens.

The advantages of the *benchmark-safety factor* (BM-SF) approach proposed by Crump are similar to the other curve-fitting methods. All of the dose-response data is used in defining the dose-response curve, and the shape of the dose-response curve and its statistical confidence limits are used directly in the estimation of the BD and the derivation of the ADI. The BM-SF approach also has greater flexibility than the other mathematical approaches because either quantal or continuous data can be analyzed, and a variety of possible curve-fitting models are available.

5. Summary and Critique

The various mathematical procedures for ADI estimation proposed by Gaylor, Dourson and Crump have not yet become widely accepted for regulatory standard-setting (Dourson and Stara, 1983). However, these procedures serve as alternative methods for comparing and validating ADIs derived from the traditional NOEL-SF approach. Unfortunately, such curve-fitting models often lack sufficient data points to provide a reliable estimate of the critical dose-response parameters (e.g. slope) in many acute or chronic animal studies (Dourson *et al.* 1986). In such situations, the regulatory agencies must continue to rely on the conventional NOEL-SF approach; otherwise many toxic chemicals currently in commercial use could not be regulated. However, Crump (1986) argues that any ADI derived from experiments with insufficient data points to apply basic statistical methods should in any case be viewed with caution. He suggests that the use of the BD model and other curve-fitting techniques would encourage better experimental design and more complete analysis of toxicological data, and promote the wider use of quantitative risk assessment methods for threshold toxicants.

Conversely, it has also been argued that problems related to limited numbers of animals and absence of estimation of slope are less critical than believed by some statisticians (Lu,

1985). For example, Lu cites toxicological studies where a 100-fold increase in the number of animals in a typical dose-response experiment has been shown to reduce estimates of benchmark risk (e.g. the VSD for ethylene thiourea) by 24-78% only, i.e., less than one order of magnitude (Lu, 1985). These differences are relatively insignificant compared to the much larger biological variability encountered in typical toxicological studies. Lu also contends that part of the safety factor for interspecies extrapolation is intended to account for uncertainties in dose-response arising from differences in group size.

Lu recommends that more resources should instead be allocated to additional experiments to uncover previously undetected pathological endpoints, such as delayed neurotoxicity or teratogenic effects. He concludes:

> *It is doubtful that any increase in the number of animals used in routine long-term studies (sic) to increase sample size and possible determination of the slope of the dose-response curve could compensate for the lack of these and other relevant tests.*
>
> Lu (1988)

Therefore, issues pertaining to the selection of the most appropriate studies and the most relevant critical endpoints for ADI determination may be at least as important as model selection and biostatistical techniques (Dourson *et al.*, 1986). Individual modeling of separate effects within a toxicological study may be necessary, with the selection of the resulting ADI values based on the most critical biological endpoint, expressed in terms of severity or frequency of occurrence. Comparison of ADIs derived from conventional NOEL-SF methods and ADIs derived by curve-fitting procedures such as benchmark dose may also assist in the selection and validation of safety factors.

C. Severity of Effects

1. Severity Scales

In order to derive an exposure standard in a realistic manner, severe or irreversible endpoints must be accorded a greater weight than mild or reversible endpoints, and the most severe effects should therefore be evaluated prior to consideration of less severe effects. Also, the magnitude of safety factors for an exposure standard are often more stringent for agents causing severe effects than for those causing mild effects. In the past, the severity of biological endpoints has played a significant but poorly defined role in the determination of NOEL values and safety factors, and has consequently influenced the derivation of ADI exposure standards. However, at present there are no standardized criteria or guidelines in the traditional NOEL-SF methodology for the incorporation of severity measures into the formal derivation of the ADI. The 1988 EPA guidelines have deferred any specific recommendations concerning severity until such methods are better validated (Barnes and Dourson, 1988).

Table 3.1
EPA Rating Values (RVe) Used to Derive Reportable Quantities Based on Chronic Toxicity

Rating	Effect
1	Enzyme induction or other biochemical change with no pathologic changes and no change in organ weight.
2	Enzyme induction and subcellular proliferation or other changes in organelles, but no other apparent effects.
3	Hyperplasia, hypertrophy, or atrophy, but no changes in organ weights.
4	Hyperplasia, hypertrophy, or atrophy, but changes in organ weights.
5	Reversible cellular changes: cloudy swelling, hydropic change, or fatty changes.
6	Necrosis or metaplasia with no apparent behavioural, sensory, or physiologic changes.
7	Necrosis, atrophy, hypertrophy, or metaplasia with a detectable decrement of organ functions. Any neuropathy with a measurable change in behavioural, sensory, or physiologic activity.
8	Necrosis, atrophy, hypertrophy, or metaplasia with definite organ dysfunction. Any neuropathy with gross changes in behaviour, sensory, or motor performance. Any decrease in reproductive capacity. Any evidence of fetotoxicity.
9	Pronounced pathologic changes with severe organ dysfunction. Any neuropathy with loss of behavioural or motor control or loss of sensory ability. Reproductive dysfunction. Any teratogenic effect with maternal toxicity.
10	Death or pronounced life shortening. Any teratogenic effect without signs of maternal toxicity.

Source: Hartung and Durkin (1986).

Since the listing of endpoints provides only a qualitative description of toxicity, there is no obvious quantitative method for numerically rating the severity of toxic effects on an objective scale. Nonetheless, based on clinical experience and scientific consensus, several severity rating systems have been developed that rank-order toxic endpoints on a semi-quantitative scale. The EPA has developed a severity scale based on chronic toxicity, which has been used as the basis for determining reportable quantities (RQs) of chemicals in toxic waste dumps (Hartung and Durkin, 1986; Stara *et al.*, 1987). The *Rating Values for Effects* (RVe) scale developed by EPA is based primarily on objective severity ratings, but is supplemented by selected subjective concerns about health outcomes (Table 3.1). The scale provides a rank-ordered progression of adverse health effects from one (mild) to 10 (severe). The first six ratings describe primarily sub-clinical cellular changes that do not give rise to any visible impairment in physiological function. Ratings 7, 8, and 9 represent three classes of organ dysfunction of increasing severity: detectable, definite, and severe. Rating 10 is reserved for lethal endpoints.

Some effects such as varying levels of teratogenic or neurological effects were included within specific ratings due to high perceived levels of societal concern.

2. NOEL Estimated by Dose-Severity Diagrams

To assess the combined dose/severity patterns for systemic toxicants and to assist in the estimation of ADI, the EPA has recently devised a system that combines the *Rating Values for Effects* (RV_e) with a separate system for rating the dose level (potency) for a given chemical (Stara *et al.*, 1987). The RV_e scale is divided into several levels of effects in increasing order of severity:

- NOAEL
  no-adverse-effect level $RV_e = 1\text{-}3$ or 4
- AEL (or LOAEL)
  adverse-effect level $RV_e = 5\text{-}7$
- FEL
  frank-effect level $RV_e = 8\text{-}10$

These effects are more fully delineated in the descriptive EPA scale of toxic effects (Table 3.1) developed by Hartung and Durkin (1986).

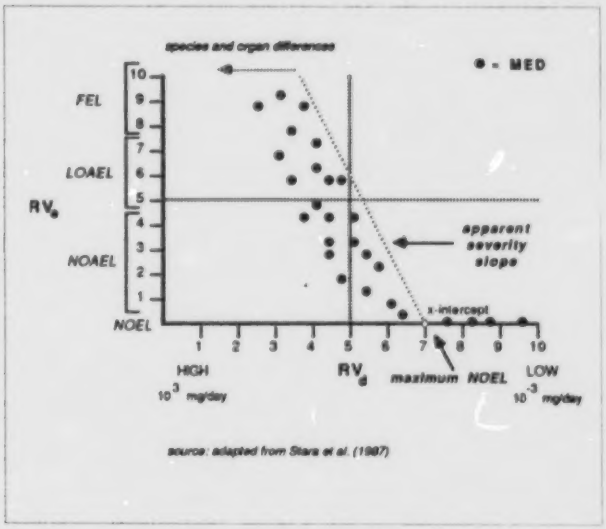
Next, a *Rating Value for Dose Level* (RV_d) is defined according to its *minimum effective dose* (MED). The MED is a general measure of the toxic potency of a given chemical in terms of dose in milligrams per day required to produce a detectable response at a specified level of severity. The MED is transformed to values ranging from one to 10 on a logarithmic scale, defined as:

$$RV_d = -1.5 \log MED + 5.5$$

for a range of log MED values between -3 and +3. Beyond this range, the RV_d is set at one when the MED is greater than 10^3 mg/day, and the RV_d is set at 10 when the MED is less than 10^{-3} mg/day. Since potency is inversely proportional to the MED, substances exhibiting an effect at low dose will have a high rating on the RV_d scale.

A graphical representation of the dose-effect relationship for a chemical can then be plotted as a two-dimensional scattergram of the RV_e and RV_d values for various biological endpoints in different organs (Figure 3.4). Each data point in the RV_d versus RV_e diagram represents the MED value for a particular biological endpoint in a specific organ or species. As the dose increases (i.e. the RV_d decreases), the effect ratings increase, moving upwards from mild NOAEL effects to moderate LOAEL effects, and finally to serious FEL effects. A variety of different biological endpoints are represented by the points in the diagram, with multiple effects appearing as the dose increases.

Figure 3.4
NOEL Estimation by Dose-Severity Diagrams Using RV_e versus RV_d Plot



For a given RV_e level, the point furthest to the *right* represents the occurrence of an effect in the most sensitive organ or species. Consequently, the line drawn to the right of all the RV_d vs RV_e data points defines the minimal dose required to produce an effect in the most sensitive organ or species at various levels of severity. The slope of this line, which is termed the *apparent severity slope*, can be extrapolated downward to intersect the x-axis immediately to the left of several NOELs for different organs or species. The x-intercept of the apparent severity slope is thus defined as the *maximum* NOEL, which is then considered as the NOEL for the chemical, taking into account effects occurring at all levels of severity and for multiple biological endpoints in different organs and species.

The RV_d vs RV_e method possess several advantages over conventional dose-response plots in the derivation of ADIs. The RV_d vs RV_e plot may be treated as a quantitative dose-effect relationship because the RV_e and RV_d values are intended to be approximations of an interval scale, rather than merely a rank-ordered scale of dose and severity of effect. For the derivation of the NOEL and ADI, it is therefore possible to apply parametric statistical methods to the estimation of the apparent severity slope and the x-intercept, including all of the available data for different endpoints at various levels of severity obtained from several chronic and subchronic experiments. In addition, the method is especially useful when there is no suitable NOAEL or LOAEL values available from experimental data. In this instance, data based on FELs and NOELs could be used to estimate the maximum NOEL or an appropriate NOAEL (e.g. a dose associated with an RV_e of one or two).

The method described above has been developed by EPA scientists in order to account for severity of effect in estimation of the NOEL. Although it has not yet been formally incorporated into the guidelines for NOEL-SF determination of ADIs, the method has been applied to assess the dose/severity relationships for over 250 systemic toxicants, and it may represent an effective future alternative to conventional approaches to ADI determination.

4. Safety Factor Approaches for Carcinogens

A. General Background

In the past two decades, conventional approaches to risk assessment for carcinogens have been based on the supposition that all carcinogens are genotoxic, acting by a stochastic process that involves random damage to the cell DNA. Consequently, most carcinogen risk assessment models adhere to the assumption of non-threshold linearity, which states that there is no threshold for the induction of cancer even at the lowest possible exposures, and that the probability of cancer induction at low levels of carcinogen exposure is a linear function of dose. Therefore, quantitative risk assessment (QRA) methods for chemical carcinogens have employed mathematical low-dose extrapolation models for the estimation of carcinogenic risk, based on the concept of the virtually safe dose (VSD), a benchmark dose corresponding to an "acceptable" lifetime cancer risk of 10^{-6} or 10^{-5}.

No explicit safety factors are included in the QRA methodology. However, the calculated exposure standards are extremely stringent, due to the selection of the most conservative extrapolation models (linear 1-hit or linearized multistage) and the use of statistical upper bounds (95% confidence limits) of the dose-response curve for the estimation of carcinogenic risk. The inherent conservatism of the assumptions included in QRA modeling can be considered equivalent to the safety factors used in the NOEL-SF methodology.

The NOEL-SF approach for deriving exposure limits for carcinogenic agents has been generally rejected by many regulatory agencies, especially by the U.S. EPA, which has been the strongest advocate of QRA methods for carcinogen regulation (Anderson, 1985; EPA, 1986b). Similarly, most consensus guidelines produced by various advisory bodies in the United States have consistently supported the non-threshold linearity assumption and its application to carcinogen risk assessment (IRLG, 1979; OSTP, 1986). In recent years, however, EPA and other U.S. regulatory agencies have reconsidered the potential use of NOEL-SF methods in special cases where the assumption of non-threshold linearity may not be relevant.

In practice, revised regulatory standards for several important carcinogenic substances have recently been proposed by EPA on the basis of mechanistic or pharmacokinetic considerations, which allow the use of threshold models that include a safety factor. Similar approaches have also been used in certain cases by other regulatory agencies in Canada and other countries.

B. Traditional Safety Factors for Carcinogens

The traditional NOEL-SF approach has occasionally been employed for the derivation of exposure limits for carcinogens, with the inclusion of added safety factors to account for the added uncertainties arising from the unique properties of carcinogenic agents. Weil (1972) initially proposed a 5000-fold safety factor for carcinogens. This value was a composite of the usual 100-fold safety factor, with an additional 5-fold factor to reflect the likelihood that the lowest applied dose represented a LOEL rather than a NOEL, plus a 10-fold factor to account for the general irreversibility of the carcinogenic process.

Health and Welfare Canada has on several occasions derived regulatory exposure limits for carcinogens based on the NOEL-SF approach on a case-by-case basis (Somers, 1981, 1984). In 1981, the allowable concentration limit of TCDD (dioxin) in commercial fish was set at 20 ppt, calculated using a 2000-fold safety factor applied to the LOEL value from rodent dose-response experiments. The rationale for employing an NOEL-SF approach was based on the lack of evidence for a genotoxic mechanism for TCDD in mutagenicity tests, its non-reactivity with DNA, and its mechanism of action involving enzyme induction and possibly other epigenetic regulatory effects. Similarly, the indoor air guideline for formaldehyde gas was set at 0.1 ppm, based on a 20-fold reduction in the existing standard for workplace air.

Clearly these decisions do not represent a rigorous approach to carcinogen risk assessment, but they do allow for the simplified determination of exposure limits for carcinogens that exhibit threshold-like dose-response behavior. However, it is unclear why the same approach was not used for establishing exposure limits for nitrilotriacetic acid (NTA) in drinking water. NTA acts as a kidney and bladder carcinogen

in rodents only at extremely high doses that induce a cytotoxic response in the urinary tract of test animals. The cytotoxic response was interpreted as a precursor to the induction of the urinary tract tumors. The use of NTA as an additive in laundry detergent was permitted in Canada (but not in the United States) on the grounds that the much lower doses of NTA typically encountered in drinking water would be unlikely to induce the cytotoxic effects, and therefore cancer induction was precluded. However, the risk assessment performed by Health and Welfare utilized a conventional low-dose extrapolation model, although there was no evidence that NTA was genotoxic or reactive with DNA.

More recently, the American Conference of Government Industrial Hygienists (ACGIH) has produced a policy paper describing the revised ACGIH guidelines for carcinogens (ACGIH, 1986), which reaffirms the long-standing ACGIH tradition of applying the safety factor approach for the risk assessment of carcinogens. According to these guidelines, the ACGIH classification procedure for carcinogens defines two categories of carcinogens:

A1: *confirmed human carcinogens*
based on epidemiological evidence in humans

A2: *suspected human carcinogens*
based either on limited epidemiological evidence, or carcinogenicity in test animals

For setting occupational exposure standards expressed as threshold limit values (TLVs), ACGIH advocates that A1 carcinogens should not be assigned a TLV, based on lack of evidence of an acceptable exposure level. Essentially this represents a zero-tolerance approach, where the substance is banned or removed from contact with the worker by protective measures. For A2 carcinogens, an ADI approach is advocated, where the exposure limit is determined by dividing the NOEL by an arbitrary safety factor of 100 or 1000 to allow an adequate margin of safety. No method is provided for the specification of an appropriate safety factor.

The stated rationale for this approach is based on the biological argument that body defenses can neutralize small doses of carcinogens, and on the methodological argument that mathematical models for low-dose extrapolation are too controversial and subjective for use in standard-setting. The traditional threshold model of toxic action is therefore held to be applicable to carcinogens, although this simplistic assumption is not in accordance with the prevailing views of most other scientific and regulatory bodies.

C. Weight-of-Evidence Classification Approach

The 1986 EPA *Guidelines for Carcinogen Risk Assessment* (EPA, 1986b) include a classification system for carcinogens that delineates the criteria for assigning various levels of "weight-of-evidence" (WOE) to specific carcinogens (Table 4.1). A ranking system consisting of five groups (Group A to Group E) describes the degree of confidence that a given chemical agent represents a genuine human carcinogen, based on epidemiological and chronic bioassay data.

Table 4.1
EPA Guidelines Weight-of-Evidence Scheme for Carcinogens

Group A -	*Human Carcinogen:* sufficient evidence in humans
Group B -	*Probable Human Carcinogen:* limited evidence in humans or no evidence in humans but sufficient evidence in animals
Group C -	*Possible Human Carcinogen:* limited or equivocal evidence in animals in the absence of human data
Group D -	*Not Classifiable as to Human Carcinogenicity:* inadequate or no data available
Groupe E -	*No Evidence of Carcinogenicity for Humans:* negative eivdence in at least two species

Source: EPA (1986b).

In addition to assisting in the identification of potential carcinogens, the EPA ranking scheme has also been used in the selection of appropriate methodological approaches for the derivation of EPA recommended standards for drinking water contaminants — including synthetic organic chemicals (SOCs), volatile organic chemicals (VOCs), and inorganic chemicals (EPA, 1985). Such a recommended standard is currently referred to as "maximum contaminant level goal" (MCLG), but was previously termed "recommended maximum contaminant limit" (RMCL).

The EPA Office of Drinking Water currently uses a three-category approach for setting MCLGs (Table 4.2), on the basis of the EPA weight-of-evidence classification for carcinogens (Cotruvo, 1988). Category I ("probable carcinogens") consists of Group A and Group B carcinogens where the weight-of-evidence is high. The MCLG for any Category I substance in water is zero concentration, as an idealized value which is not necessarily attainable in practice. MCGLs are thus considered as "aspiration" goals, not standards.

Category II ("equivocal carcinogens") consists of Group C substances which have a lower weight-of-evidence value. EPA allows derivation of MCGLs for Category II substances either by calculating the VSD that produces a lifetime excess cancer risk of 10^{-5} or 10^{-6} using a conservative low-dose extrapolation model, or by calculating an ADI (RfD) value with added safety factors that allow for an "adequate margin of safety". Any MCGLs calculated by the NOEL-SF method are based on non-carcinogenic endpoints. The NOEL-SF method is the preferred mode of derivation where appropriate non-carcinogenic data is available. The ADI is calculated using the normal safety factor of 100; then an added safety factor of up to 10 is applied to account for additional uncertainties on a case-by-case basis.

Table 4.2
EPA Three-Category Approach to Deriving Maximum Contaminant Level Goals (MCLGs) for Drinking Water Contaminants

Category 1 ("probable carcinogens")
group A and B substances
MCGL equals zero (as aspirational goal)

Category 2 ("equivocal carcinogens")
Group C substances
MCLG equals VSD at 10^{-5} to 10^{-6} excess cancer risk,
or
MCLG equals ADI with added safety factors
ADI = NOAEL divided by 1000
standard 100-fold SF with additional 10-fold SF for chemicals with "equivocal evidence of carcinogenicity"
or
a safety factor other than 10 if the properties of the chemical dictate an alternative factor

Category 3
Group D and E substances
treated as threshold toxicants (non-carcinogens)
ADI approach with portion allocated to drinking water

Source: Cotruvo (1988).

This methodology has been applied for the derivation of RMCLs for six SOCs in the 1985 *National Primary Drinking Water Regulations* (EPA, 1985). The six chemicals placed into Category II were: styrene; 1,2-dichloropropane; monochlorobenzene; lindane; asbestos; and arsenic. For such Category II chemicals, the ADI is normally derived by dividing the NOAEL by a safety factor of 1000, comprising the standard 100-fold safety factor with an additional 10-fold safety factor for chemicals with "equivocal evidence of carcinogenicity" (EPA, 1985). However a factor other than 10 can be applied if the properties of the chemical indicate that an alternative factor is more appropriate.

For example, the RMCL for styrene was based on chronic toxicity data because carcinogenic bioassay data was very limited and inadequate for extrapolation purposes. The ADI for styrene was calculated by dividing the NOAEL by a 1000-fold safety factor, where the extra 10-fold safety factor was included to account for a less-than-lifetime exposure in the test animals (EPA, 1985). The RMCL for styrene was then derived by dividing the ADI (7 mg/L) by an additional 10-fold safety factor to account for its potential carcinogenicity, and a 5-fold exposure factor that assumes a 20% contribution of drinking water to total intake. The final RMCL was thus calculated as the ADI divided by 50, or 0.14 mg/L.

Category III consists of Group D and E substances, where the probability of carcinogenic action is low or negligible. These chemicals are treated as threshold toxicants (non-carcinogens) using the ADI approach, with an adjustment made for the proportion of intake allocated to drinking water. For example, the solvent 1,1,1 trichloroethane (EPA Group D) was considered a Category III substance, and was regulated as a non-carcinogen under the 1987 EPA drinking water standards specified for eight VOCs (Cotruvo, 1988).

D. Mechanistic Classification Approach

Conventional methodologies for the risk assessment of carcinogens are predicated on the assumption that all carcinogens act through mutagenic damage to the cell DNA, and therefore all carcinogenic chemicals should be considered as genotoxic agents. However in recent years, an accumulating body of evidence has indicated that carcinogenesis can take place through several different modes of action, thereby allowing for the mechanistic classification of chemicals into two fundamental categories: (1) *genotoxic* carcinogens which cause somatic mutations either by direct DNA damage or by indirect interactions with other cellular genetic components, and (2) *epigenetic* carcinogens which affect cell regulation and gene expression without alteration of the DNA itself (Weisberger and Williams, 1981).

According to a somewhat simplistic interpretation of the multistage model of carcinogenesis, tumor initiators are generally held equivalent to genotoxic carcinogens, while tumor promoters, carcinogenic hormones, cytotoxic agents, and peroxisome proliferators are classified as epigenetic carcinogens. One of the principal consequences of such a mechanistic distinction relates to the dose-response behavior of the chemical in question. Genotoxic carcinogens are assumed to follow a linear non-threshold dose-response relationship and exhibit non-zero risk at low doses. In contrast, epigenetic carcinogens are believed to possess a threshold of cellular toxicity, below which the risk for cancer induction is zero (Weisberger and Williams, 1981):

> *Most epigenetically active agents are active only at high sustained doses and, up to a certain point, the effects they induce are reversible. Thus, these types of agents may represent only quantitative hazards to humans, and it may be possible to formulate safe levels of exposure after appropriate toxicological dose-response studies are conducted.*
>
> Weisberger and Williams (1981)

Since the formulation of the genotoxic/epigenetic distinction in the early 1980's, several regulatory agencies have considered introducing a mechanistic classification approach for the regulation of carcinogens. The potential regulatory significance of a mechanistic approach to carcinogen risk assessment is considerable. In principle, epigenetic carcinogens can be considered as adaptable to the NOEL-SF approach for regulatory standard-setting, providing at least a nominal assurance of "absolute safety" for this class of carcinogen. More important, risk estimates determined using the NOEL-SF approach are often several orders of magnitude lower than those produced by the conservative low-dose extrapolation models (Perera, 1984). Hence the exposure limits for epigenetic carcinogens calculated using safety factors may be several thousand-fold less stringent than exposure limits

derived from low-dose extrapolation models. However, the introduction of extra safety factors to account for the uncertainties of NOEL determination in carcinogen bioassay studies, and to reflect the irreversibility of carcinogenic action, may greatly diminish this supposed advantage (EPA, 1985).

The greatest scientific and political debates over the application of the genotoxic/epigenetic distinction and the use of the NOEL-SF approach, have taken place in the United States, especially within the EPA (Marshall, 1982; Perera and Petito, 1982; Ashford *et al.*, 1983; Perera, 1984). The original 1976 EPA Cancer Policy proposed a rigorous adherence to the linear non-threshold model for the risk assessment of all carcinogens. In 1979, this policy was further substantiated by the report of the Interagency Regulatory Liaison Group (IRLG, 1979). The IRLG guidelines for carcinogen risk assessment strongly supported the EPA position by stating that a no-effect threshold could NOT be assumed for carcinogenic substances. Only a cautious acknowledgement of possible mechanistic interpretations of test data was allowed in exceptional cases.

However, in 1982 EPA administrators attempted to put forward a draft proposal containing policy revisions for the risk assessment practices of the agency, as an "Addendum" to the 1976 EPA Cancer Policy (Marshall, 1982; Perera and Petito, 1982). The addendum recommended a new genotoxicity classification system based on positive results in two different *in vitro* mutagenicity tests (one microbial and one eukaryotic), or on a single positive mammalian *in vivo* test. The proposal was conceptually based on the mechanistic classification system of Weisberger and Williams (1981), but it contained a more limited number of tests with no specifications for a battery of mutagenicity tests.

The most important regulatory application of the new EPA policy was a proposal for the regulation of formaldehyde (FA) gas as an epigenetic carcinogen. The 1982 Todhunter memorandum recommended that FA should be regulated by EPA as a threshold carcinogen based on its observed cytotoxicity in the nasal epithelium of test animals and humans (Perera and Petito, 1982). The proposal for a mechanistic FA exposure standard cited the Weisberger and Williams (1981) paper as the rationale for the epigenetic/genotoxic distinction. However, the Todhunter memorandum was severely criticized on scientific and procedural grounds both within the EPA and by its advisory bodies, and was subsequently withdrawn (Perera and Petito, 1982; Ashford *et al.*, 1983). The mechanistic cancer policy of EPA was later abandoned as a consequence of inadequate scientific consensus and administrative turmoil within the agency (the "Sewergate" scandal).

Later attempts to revive the mechanistic approach within the U.S. regulatory system were met with considerable suspicion in the scientific community. During the initial drafting of the report of the Office of Science Technology Policy on chemical carcinogenesis (OSTP, 1986), the mechanistic approach was again proposed, but subsequently was toned down in response to criticism by many scientists both in the regulatory and research communities (Rushefsky, 1985). Similar considerations affected the drafting of the 1986 *EPA Guidelines for Carcinogen Risk Assessment* (EPA, 1986b).

Among the most serious criticisms relating to the mechanistic approach is the stipulation that a toxicological threshold must exist for the action of promoters and other epigenetic carcinogens. This premise has been poorly validated, because adequate dose-response studies have not been performed, and existing promoter studies are generally not designed with dose-response analysis in mind (OSTP, 1986). At a more fundamental level, Perera (1984) has criticized the concept of a genotoxic/epigenetic distinction for carcinogens on the grounds that promoters may act through genotoxic mechanisms. Her arguments can be summarized as follows:

- no clear cut or rigorous biological distinction exists between "genotoxic" and "epigenetic" carcinogens
- one cannot assume the most single plausible or significant mechanism is necessarily genotoxic or that genotoxic/epigenetic are mutually exclusive
- it is simplistic to identify genotoxic activity as equivalent to initiation and epigenetic activity to promotion
- dose-response curves are not necessarily different for epigenetic carcinogens, nor can thresholds be assumed; therefore epigenetic carcinogens cannot be considered as "safer" than genotoxic carcinogens
- the tumor inducing potency of many promoters is very great (e.g. TPA, TCDD), and can be as potent as powerful classical carcinogens

In addition, the OSTP report notes that population thresholds for an epigenetic carcinogen may vary even if a toxicological threshold can be determined in test animals (OSTP, 1986). However, this objection could also be raised for any conventional toxicant, and can be accounted for by a suitable population safety factor. At present, the arguments for and against the mechanistic approach to risk assessment can be summarized as follows:

> *Those opposed to this perspective argue that negative studies on the current battery of genotoxicity (tests) by these so-called epigenetic carcinogens do not unequivocally prove that the agent is not acting on the DNA or that it is acting by a threshold response. They argue that the lack of sensitivity of current tests to discern the capacity of all carcinogens to initiate DNA should not be an excuse to create a new belief system which clearly does not err on the side of safety. The proponents of the epigenetic position counter by saying that it is clearly impossible to prove a negative and that because of philosophical orientation, such scientists would still not be satisfied if a dozen or two of the new genotoxicity tests were negative.*

This issue may go unresolved, but it remains important with major economic and public health implications. If the epigenetic perspective were sustained, then the issue of what would be an acceptable safety factor for carcinogens would emerge.

Calabrese (1983)

The recent EPA *Guidelines for Carcinogen Risk Assessment* (EPA, 1986b) are cautiously permissive with respect to mechanistic approaches to risk assessment, allowing the use of safety factors under conditions where the evidence of promoter action or another epigenetic mechanism is particularly strong:

When pharmacokinetic or metabolic data are available, or when other substantial evidence on the mechanistic aspects of the carcinogenesis process exists, a low-dose extrapolation model other than the linearized multistage procedure might be considered more appropriate on biological grounds.

EPA (1986b)

The use of safety factor approaches for carcinogen risk assessment is best illustrated by recent revisions to the water exposure standard for dioxin (TCDD), undertaken both by the EPA in the United States and by the Ontario Ministry of the Environment (OME) in Canada. The use of pharmacokinetic approaches is illustrated by the use of PB-PK modeling by EPA for the solvent methylene chloride. These methodologies will be examined in detail in the following sections.

E. Mechanistic Approaches for the Risk Assessment of Dioxin

1. EPA Methodology

Significant changes in the EPA risk assessment methodology for carcinogenic agents have recently been proposed in deriving revised exposure limits for dioxin — specifically 2,3,7,8-tetrachlorodibenzo-p-dioxin, or TCDD (EPA, 1987; Shu *et al.*, 1987; Gough, 1988; Finkel, 1988). Both the existing and revised EPA risk estimates are based on the same liver tumor dose-response data for TCDD in rats (Kociba *et al.*, 1978). In these studies, a frank liver tumor response was seen at a daily intake level of 0.1 µg/kg/day, a slight non-tumor response in the form of liver enzyme induction at 0.01 µg/kg/day, and no observed response at 0.001 µg/kg/day. From this data, application of a conventional linearized multistage extrapolation model for TCDD produces an extremely stringent daily intake limit of 0.006 picograms (pg) per kilogram (kg) body weight per day (EPA, 1985), where 1 pg equals 10^{-6} micrograms (µg).

Such a restrictive intake limit for TCDD is roughly equivalent to a daily intake of 10^7 molecules of TCDD per kilogram body weight, a quantity which is unlikely to exert any significant biological or carcinogenic effect, unless the agent is a potent genotoxic agent. In fact, TCDD exhibits virtually no genotoxic or tumor initiating activity in extensive batteries of *in vitro* and *in vivo* mutagenicity tests, and shows no evidence of chemical interaction with DNA (Shu *et al.*, 1987).

Conversely, TCDD has been shown to act as a potent tumor promoter in several systems, including rat liver previously initiated with diethylnitrosamine and in mouse skin initiated with polycyclic aromatic hydrocarbons (Gough, 1988). These findings suggest that TCDD may induce cancer solely by promotion of cells that have previously been exposed to environmental initiator substances, or which are genetically predisposed to spontaneous tumor initiation. Moreover, the biological activities of TCDD, including its cytotoxic and tumor promoting properties are entirely dependent on its ability to bind to the cytosolic *Ah* receptor protein found in a variety of mammalian cells. Because the extent of receptor-mediated binding is a function of TCDD concentration and receptor-ligand affinity, the dose-response kinetics of TCDD promoter action would be expected to behave in a deterministic fashion, thereby exhibiting a true threshold for promoter activity at low doses.

In view of these mechanistic considerations, fundamental changes in the EPA regulatory policy for TCDD were proposed in a 1987 draft report (EPA, 1987), which recommended that the conventional risk estimation based on the linearized multistage model should be modified by the inclusion of a safety factor approach in the risk estimation procedure. The impetus for this change originated in part from the regulatory difficulties that EPA experienced in enforcing the extremely stringent daily intake standard for TCDD using the extrapolation modeling approach, where an intake limit of 0.006 pg/kg/day was required to achieve a benchmark risk (VSD) of 10^{-6} lifetime risk. This value represented an almost impossible regulatory standard for human exposures to TCDD in fish, toxic waste effluents, soil contamination, and incinerator emissions. The revised intake limit for TCDD proposed in the EPA draft report was set at 0.1 pg/kg/day, which indicates that the estimated carcinogenic risk from TCDD exposure is 17-fold lower when the safety factor approach is included in the risk assessment.

The proposed EPA daily intake limit for TCDD is a peculiar hybrid representing the midpoint of six different risk assessments obtained both from the low-dose extrapolation and NOEL-SF approaches. (EPA, 1987; Finkel, 1988; Gough, 1988; Perera, 1988). The EPA rationale for including the NOEL-SF method is based on the possibility that TCDD could act as an epigenetic promoter in humans. Accordingly, EPA has derived an ADI for TCDD using 1000 pg/kg/day (0.001 µg/kg/day) as the NOEL for tumor promotion, and has selected a safety factor of either 100 or 1000 to produce an ADI value of 10 pg/kg/day or 1 pg/kg/day respectively. These ADI limits allow for a 1670-fold or 167-fold greater exposure to TCDD than the corresponding limits, 0.006 pg/kg/day, derived from the linearized multistage model.

The more conservative of the two ADI values derived from the NOEL-SF method (1 pg/kg/day) is then "averaged out" in combination with the daily intake value obtained from the low-dose extrapolation method (0.006 pg/kg/day) to produce a

composite daily intake limit of 0.1 pg/kg/day. Converted for a 70 kg person, the respective daily exposure limits for the NOEL-SF and the low-dose extrapolation methods are 70 pg/day and 0.42 pg/day, while the averaged value is 7 pg/day (0.1 pg/kg/day × 70 kg). The averaged value for the TCDD exposure limit appears to approximate the geometric mean of the exposure limits obtained from the NOEL-SF and low-dose extrapolation methods. The averaged value is a 17-fold greater exposure limit than the conventional limit derived from the low-dose extrapolation method.

The EPA justification for the revised TCDD intake daily limit is that the value obtained lies "near the mid-point" of a range of previous EPA estimates and of a group of lower estimates from regulatory agencies in other countries (including Canada), where the safety factor approach has been used almost exclusively (Gough, 1988). Thus the EPA risk estimate is based on a combination of threshold and non-threshold models, which in theory are mutually irreconcilable. Most surprisingly, the revised EPA risk estimates for TCDD have met with severe criticism from some quarters, and the agency has responded defensively by stating that the proposed changes in methodology represent merely an attempt to advance the state of the art in risk assessment (Finkel, 1988).

> *Instead, EPA is now on record saying that dioxin is "sort of" a promoter and that its potency is "sort of" lower than it once thought.*
>
> Finkel (1988)

On a more practical level, considering that the average daily intake of TCDD in the United States is about 1 pg/day, the acceptance of the proposed 7 pg/day daily intake limit as the official exposure standard would greatly facilitate the setting of "reasonable" exposure standards for environmental contamination by TCDD (Perera, 1988).

In contrast, the U.S. Food and Drug Administration (FDA) has traditionally employed a simplified safety factor approach for deriving tolerances for chemical contaminants in edible fish, which has generally produced higher allowable exposure limits than the corresponding EPA standards (Shu *et al.*, 1987). For TCDD, the ADI value is calculated simply as the NOEL (1000 pg/kg/day) divided by a modest safety factor of 77, to produce an ADI value of 13 pg/kg/day. The FDA limit is about 2000 times the conventional EPA limit for TCDD, and still greatly exceeds the revised EPA limit. Based on the estimated fish consumption levels at the 99th percentile, the FDA tolerance for TCDD in edible fish is set at 25 ppt.

2. OME Methodology

In the review by Shu *et al.* (1987) describing the cancer risk assessment for TCDD, the conservative low-dose extrapolation approach traditionally advocated by EPA is contrasted with a much less stringent NOEL-SF methodology utilized by the Ontario Ministry of the Environment (OME). On the basis of this methodology, Ontario and some Western European countries consider that the carcinogenic risk of TCDD is significantly lower than the conservative risk estimates produced by EPA and other U.S. regulatory agencies.

The OME methodology treats TCDD as an epigenetic promoter with threshold activity, and thus it applies a simple safety factor model for derivation of the daily intake limit. The OME risk estimation is derived using a NOEL of 1000 pg/kg/day (the same NOEL value selected by EPA in their revised method), and dividing by a safety factor of 100 to obtain a daily intake limit of 10 pg/kg/day for humans. This value is approximately 1670 times more tolerant than the existing EPA daily intake limit of 0.006 pg/kg/day.

OME has based its rationale for the use of the NOEL-SF approach on the existence of a theoretical threshold for carcinogenic activity demonstrated by promoter activity in animals, and on the empirical determination of a threshold for promotion using the NOEL value obtained from chronic bioassays. The OME position can be summarized in the following points (quoted from Shu *et al.*, 1987):

- The NOEL (0.001 µg/kg/day) identified in rodent carcinogenicity bioassay studies indicates where the threshold level for tumor production by TCDD exists.
- While TCDD has been rated as the most potent carcinogen using absolute quantities as criterion, this must be viewed in the context that its carcinogenic properties are expressed at concentrations only two to three orders of magnitude below the LD_{50} range.
- Mutagenicity studies, judged on a battery of short-term tests, indicate that TCDD is not a mutagen in the classical sense. The lack of evidence to suggest that TCDD or its metabolites can directly alter DNA physically or chemically also supports this conclusion.
- From the data in the above section, it can be concluded that PCDDs and PCDFs, especially TCDD, can produce tumors in rodents by an indirect mechanism. A threshold dose exists, as indicated by NOELs from long-term studies.

Hence the OME risk assessment regards TCDD as a carcinogenic hazard only at relatively high doses approaching the levels that induce frank toxicity, where promoter effects become manifest by the induction of precursor lesions or malignancies in the target organs. At doses below the NOEL, where no such cellular or organ toxicity is detectable, the agent is deemed to be harmless due to the presence of a true toxicity threshold.

Conversely, OME has explicitly rejected the use of low-dose extrapolation models to obtain VSD risk estimates. OME cites several weaknesses in quantitative risk assessment by mathematical models that renders the methodology unacceptable for setting standards:

- There are many types of mathematical models and they produce very different risk estimates from the same biological data.
- Current models assume only direct action by the chemical; current models do not incorporate indirect models of action (e.g. tumor promoters).

- Current models extrapolate probabilities from measurements made in the 10^{-1} to 10^{-2} range down to 10^{-5}, 10^{-6}, or 10^{-8}, i.e., well beyond the realm of biological certainty.
- Use of these risk-analysis models, in this instance, should therefore be more to indicate the potential range of safe doses rather than to form the basis of a standard.
- The non-linear and sex-specific nature of the rodent bioassay data used and the presence of dose-related primary liver damage at treatment levels causing hepatocellular neoplastic change suggest that these risk estimates for cancer induction may be confounded by direct tissue damage. Lack of knowledge of the mode of action of 2,3,7,8-TCDD also precludes selection of a specific risk-analysis model.

None of these arguments are particularly novel or insightful, but the points raised do represent a useful example of the reasoning that leads many regulatory agencies outside of the United States to select the NOEL-SF approach for setting standards for those carcinogenic agents that exhibit threshold-like behavior.

Finally, OME defends its selection of a modest 100-fold safety factor for TCDD, by articulating the following points:

- The 100-fold safety factor is a practical means to handle the uncertainties in extrapolating from animals to humans. It includes a factor of 10 to extrapolate from animals to humans assuming that animals are less sensitive than humans and another factor of 10 to account for differential sensitivities within the human population.
- This factor incorporates a number of considerations to account for uncertainty in extrapolating from animal data to humans, particularly an allowance in case humans are more sensitive than the animal species tested.
- Since acute toxicity and long-term animal studies are available, and since the short-term mutagenicity studies and human epidemiology studies are generally negative, a safety- factor of 100 is recommended.
- The NOEL of 0.001 µg/kg/day for 2,3,7,8-TCDD, determined in the three-generation reproductive study of Murray *et al.* (1979) and the two-year oncology study of Kociba *et al.* (1978) (both using rats), is recommended as a prudent basis for developing a maximum allowable daily intake for human PCDD and PCDF intake.

Therefore, OME has elected to apply only the traditional 100-fold safety factor, disregarding any additional considerations concerning the carcinogenic nature of TCDD and the attendant uncertainties that are associated with analyzing the dose-response behavior of carcinogens.

In comparison, several other countries that have used the NOEL-SF approach for TCDD regulation have chosen to introduce additional safety factors that result in sizeable variations in ADI values (Shu *et al.*, 1987; Gough, 1988). In the Netherlands, an added safety factor of 2.5 has been included, to account for the possibility that unobserved marginal effects could occur in the liver of female rats, so that the apparent NOEL would actually represent a LOEL. This adjustment gives an overall safety factor of 250, which produces an ADI of 4 pg/kg/day. Similarly, the Federal Environmental Agency of the Federal Republic of Germany has specified a safety factor that ranges from 100- to 1000-fold, resulting in an ADI ranging between one to 10 pg/kg/day. The rationale states that TCDD should be accorded a higher safety factor than the usual 100-fold factor because it is a chemical contaminant.

F. PB-PK Modeling

For carcinogen risk assessment using low-dose extrapolation, it is generally assumed that the probability of tumor induction is a direct function of the applied dose of the test agent. However, most carcinogenic chemicals undergo extensive pharmacokinetic processing in the body, including metabolism via competing pathways of bioactivation and detoxication. Hence there is often a complex relationship between the *applied* (external) dose of a carcinogenic agent and the *effective* (internal) dose at critical target organs. It is the effective dose that will directly determine the extent of tumor induction at the cellular level. Many pharmacokinetic processes are saturable at high doses, especially if they involve enzyme-catalyzed metabolic steps. Consequently, the dose-response curve for carcinogens will often exhibit non-linear dose-dependent kinetics, where the mathematical function describing the dose-response relationship is different at varying dose levels (Hoel *et al.*, 1984).

Physiologically-based pharmacokinetic (PB-PK) modeling is a technique that allows for the calculation of the internal dose of a test chemical, by applying a computer simulation model to integrate information on the applied dose, the physiological compartments which the chemical occupies, and its specific metabolic pathways (Andersen *et al.*, 1987; Watanabe *et al.*, 1988). PB-PK modeling constructs a mathematical description of the fate of chemicals in the body, and provides estimates of the concentration of the activated carcinogen at the target site. PB-PK modeling does not make any assumptions regarding the mechanistic aspects of carcinogenicity, it merely provides a more accurate method for estimating internal dose (Andersen *et al.*, 1987).

Initially, it is necessary to identify the active carcinogenic species and the significant metabolic pathways that lead to the formation of proximate and ultimate carcinogens within the cell. Computer simulation models of pharmacokinetic processes require extensive physiological data, both for the test animals and for humans. Pharmacokinetic model parameters must be determined in both species in order to quantify the processes of absorption and distribution, bioactivation and

detoxication, and excretion. In addition, the kinetics of binding to cellular macromolecules and repair or fixation of cellular lesions must also be modeled.

Unfortunately, except for a few well-characterized organic solvents and monomers, the experimental data related to these parameters is often incomplete, especially for humans (Perera, 1988). The greatest problem in PB-PK modeling is that most of the pharmacokinetic parameter values for humans are based on short-term clinical measurements on a limited sample of individuals, which do not accurately reflect the wide interindividual variation in human responses. However, a limited number of assumptions and approximations can be introduced into a PB-PK model from other data sources without seriously compromising the accuracy of the model. For example, poorly defined human pharmacokinetic rate constants can be replaced by animal metabolic rate constants with suitable adjustment factors. The validity of the modeling predictions can subsequently be assessed by comparing the predicted and observed levels of the parent chemical and its bioactivated metabolites in the blood and target tissues of test animals (Andersen *et al.*, 1987).

The regulatory impact of PB-PK modeling has recently become apparent in the proposed risk assessment by EPA of two common organic solvents: methylene chloride (dichloromethane, DCM) and 1,1,1-trichloroethane (TCA). In both cases, the PB-PK approach to risk assessment has produced risk estimates that are significantly lower than those obtained by conventional techniques. Given the widespread usage and economic importance of these chemicals and other chlorinated aliphatic solvents, the PB-PK method has attracted intense scrutiny by regulatory agencies for setting revised air and water quality standards.

Methylene chloride (DCM) is a potential carcinogen, which induces lung and liver tumors in mice at high dose levels. For humans exposed to low levels of DCM, PB-PK modeling predicts that the internal dose of carcinogenic DCM metabolites at the target organs is 140- to 170-fold lower for inhalation or 50- to 210-fold lower for drinking water ingestion than the equivalent dose calculated by the conventional low-dose extrapolation method (Andersen *et al.*, 1987). Based on the unsubstantiated premise that all species will be equally sensitive to a given target tissue dose, the PB-PK model indicates that previous risk assessments derived from rodent bioassays may have greatly overestimated the carcinogenic risk to humans exposed to low concentrations of DCM.

The PB-PK risk assessment for DCM does not require the inclusion of any allometric scaling factor to account for species size differences, nor is there any need for an interspecies safety factor, because the model itself accounts for the relevant differences between animal and human pharmacokinetic parameters. Thus PB-PK modeling may permit the reduction or removal of the interspecies safety factor, although the requirement for a safety factor to account for human interindividual variability still remains. While several assumptions and uncertainties are included in the model, PB-PK risk assessments should still be considered more reliable than conventional risk assessments that do not account for pharmacokinetic processes at all.

PB-PK modeling has also been used for the risk assessment of 1,1,1-trichloroethane (TCA), a solvent which produces moderate reversible histological alterations in the liver of rodents (Watanabe *et al.*, 1988). Conventional risk assessments based on a series of "black box" animal models have suggested that TCA could accumulate to a high body burden during chronic exposure, and therefore would represent a significant health risk. In contrast, PB-BK modeling predicts that for humans consuming 2 litres/day of drinking water containing 0.3 ppm TCA, the chemical will be non-accumulative within the body and reach a steady-state level within a few days. The calculated steady-state body burden for TCA in humans will be only 4×10^{-3} mg/kg, which is only 2.7 times higher than the one-day consumption level. This level represents a body burden 5000 times lower than the level inducing minimal effects in test animals. Therefore it appears that there is no significant health risk from TCA contamination of water supplies at typical human exposure levels.

In future, PB-PK modeling may become more widely applied for risk assessment of other economically important chemicals, such as trichloroethylene and tetrachloroethylene, provided that adequate pharmacokinetic data is available. The use of this technique promises to reduce the uncertainties inherent in the low-dose extrapolation for carcinogens, and also eliminates the interspecies conversion factor for all types of chemicals.

> *The continued use of arbitrary safety factors constitutes an inflexible bias assuming humans are always more sensitive than the most sensitive animal species tested. This is likely not the case for all materials. If a selection of safety factors based on knowledge of the sensitivity in humans is employed, other factors for consideration include seriousness of the toxic effect, whether the material is a potential pharmaceutical or industrial chemical, the likely exposure regimen, and the entire scientific rigor of the data base. Finally it should be noted how well the 100-fold safety factor from the most sensitive species no-observed-effect level has served the scientific community. However, our advanced capability of understanding interspecies sensitivity and detecting more and more subtle effects may warrant a more flexible approach than was used in the past.*
>
> Watanabe *et al.* (1988)

Thus PB-PK modeling may partially remove the need for inclusion of arbitrary safety factors for setting exposure standards.

5. Safety Factor Approaches for Teratogens

A. General Background

Chemicals that cause adverse reproductive outcomes through prenatal toxicity to the embryo or fetus are referred to as teratogens or developmental toxicants. Teratogenic effects are manifested by a wide variety of abnormal developmental outcomes, including miscarriage, fetal death, structural malformations, functional deficits, reduced birthweight, or increased neonatal morbidity and mortality (IRLG, 1986). In terms of organization and evaluation of toxicity data, the risk assessment methodology for developmental toxicants has been derived largely from models previously developed for risk assessment of carcinogens. However, from a mechanistic viewpoint, developmental toxicants are generally considered as threshold toxicants that are amenable to the NOEL-SF approach of risk assessment:

> *The regulatory framework most relevant to developmental toxicity must be based on the no-effect level and a prudent safety factor, because developmental toxicity is a threshold-type phenomenon. If something can adversely affect the embryo, this does not necessitate its proscription.*
>
> Johnson (1989)

At present, the most comprehensive risk assessment methodology for teratogens has been delineated in the 1986 EPA *Guidelines for the Health Assessment of Suspect Developmental Toxicants* (EPA, 1986c). However, the quantitative procedures for assessing developmental toxicants are less well developed than for carcinogens:

> *The SAB (Science Advisory Board) Panel noted that the field of developmental toxicology is particularly weak with respect to quantitative assessment and recommended that further efforts be given to developing alternative methods for quantitative estimates of risk for developmental toxicity.*
>
> EPA (1986c)

Adequate data resources for the dose-response assessment of teratogens are only rarely available from human epidemiological studies. Therefore, for the purposes of standard-setting, bioassay tests usually supply the necessary dose-response information required for the estimation of teratogenic risk.

> *Evidence for a dose-response relationship is an important criterion in the assessment of developmental toxicity, although this may be based on limited data from standard three-dose studies.*
>
> EPA (1986c)

Unfortunately, animal teratogenicity data do not always follow a classic dose-response relationship for developmental endpoints, due to the multiplicity of potential endpoints and the effects of competing risks (e.g. malformation versus fetal death). Therefore, the use of an integrated teratogenic evaluation measure such as combined deaths and malformations is recommended, and in all cases the individual developmental endpoints must be carefully interpreted (EPA, 1986c). In addition, because the numerical measures used to quantitate various endpoints differ in their metric properties (continuous, count, and binary data), no single biostatistical procedure can be routinely recommended as the preferred procedure for all types of dose-response modeling (IRLG, 1986).

For the statistical determination of teratogenic risk at exposure levels normally encountered in occupational and environmental settings, two types of risk assessment procedures can be employed using available dose-response data from animal experiments. The two basic methods for risk estimation are: (1) the NOEL-SF approach, and (2) mathematical extrapolation models. The NOEL-SF approach presupposes the existence of a threshold for teratogenic action, below which no adverse developmental toxicity will occur. Extrapolation models do not presuppose any definitive threshold, but predict that an acceptably low level of teratogenic risk will occur at a sufficiently low dose of a test substance.

B. NOEL-SF Approach for Teratogens

According to the 1986 EPA Guidelines, the recommended dose-response methodology for the determination of maximum exposure limits is the safety factor approach (EPA, 1986c). A no-observed-effect level (NOEL) is determined based on the highest experimental dose producing no statistically significant increase in developmental abnormalities for a predefined endpoint, after subtraction of any background incidence in the concurrent control groups. The NOEL value is then divided by a safety factor to account for the uncertainties of interspecies differences and human variability, to produce a maximum exposure limit.

The EPA policy is to employ a NOEL-SF methodology based on OECD Guidelines #414 requirement for NOEL determination (OECD, 1981), with the proviso that a toxicity threshold is assumed but not proven. The NOEL value is defined for the most sensitive developmental endpoint, in the most appropriate and/or most sensitive species (EPA, 1986c). On the other hand, the 1986 EPA Guidelines also specify that no one species should be considered as the most appropriate, but rather that selection of the test species should be evaluated on a case-by-case basis. In animal tests using a minimum number of dose-levels (three), it is common to find that even the lowest experimental dose produces a significant teratogenic effect. The lowest dose is then defined as the lowest-observable-effect level (LOEL). Maximum exposure levels can be calculated by dividing the LOEL by a number that includes an additional safety factor to be determined by scientific judgement. Under conventional toxicological practice, the default value for the additional safety factor in LOEL calculations is set at 10.

Although the magnitude of the safety factor is critical in determining an exposure standard, the definition of the permissible exposure level is specified by the 1986 EPA Guidelines in an extremely vague manner as the: "*calculated exposure level believed to be unlikely to cause any toxic developmental response in humans*" (EPA, 1986c).

Thus the 1986 EPA Guidelines do not specify any predetermined value for the safety factor (uncertainty factor). Although traditional toxicological practice normally sets a safety factor of 100 for acute and chronic toxicants in adults, no quantitative criteria are specified by EPA for calculation of safety factors for teratogens. Instead, EPA states that the safety factors for teratogenic agents are variable, and should be determined on a case-by-case basis according to "scientific judgement". The EPA Guidelines do, however, specify several qualitative criteria for evaluation of safety factor values:

- *availability of NOEL versus LOEL*
  an additional safety factor is required when a LOEL is used
- *slope of the dose-response curve*
  a shallow slope extrapolates to higher risk at low dose, and requires that larger numbers of test animals be used to determine the NOEL
- *nature of interspecies differences*
  comparability and consistency of observed teratogenic endpoints in test animals versus humans will influence the safety factor
- *types of developmental effects observed*
  more severe effects require the use of a larger safety factor
- *maternal/fetal toxicity dose ratio (A/D ratio)*
  specifically fetotoxic effects require a larger safety factor

In general, the EPA Guidelines also advocate the selection of an overall safety factor that takes into account the weight-of-evidence for developmental toxicity, the reliability of the dose-response relationship, the degree of human exposure, and "*any other factors that affect the quality and precision of the assessment*" (EPA, 1986c). The role of "expert judgement" in the application of these criteria to the setting of safety factors is repeatedly emphasized by EPA.

The problem of determining appropriate values for the safety factor has been more extensively addressed in a methodology developed by the U.S. Food and Drug Administration (FDA) for drug safety evaluations (Frankos, 1985). Under this scheme, an integrated evaluation of maternal/developmental toxicity is included in the risk assessment, which classifies substances according to (1) the "*developmental specificity*" of the teratogenic endpoints in animal tests, and (2) the "*embryotoxic selectivity*" of the teratogenic dose-response relationship.

"*Developmental specificity*" indicates whether observed developmental endpoints in animal tests are specific or non-specific (Frankos, 1985; IRLG, 1986). "Specific" endpoints (Table 5.1) are revealed as structural defects and embryolethal effects (Type I), which are considered more severe and more definitively teratogenic; "non-specific" endpoints such as retarded fetal development, reduced birthweights, and decreased postnatal survival and growth (Type II) are considered less severe and less indicative of teratogenicity. Qualitatively, Type I effects should be accorded greater weight than Type II effects, and consequently this criterion may affect the magnitude of the safety factor.

"*Embryotoxic selectivity*", which is often expressed as the A/D ratio, refers to the relation between the lowest dose (A) producing overt chronic toxicity in the adult mother versus the lowest dose (D) inducing developmental toxicity in the embryo. Substances with a low A/D ratio are considered as "coeffective" agents that produce toxicity in the same dose range both in the mother and fetus, often attributable to a secondary effect on fetal development due to altered maternal homeostasis (Johnson, 1989). Generally, agents with an $A/D < 1$ have a low hazard index and do not exhibit embryotoxic selectivity, whereas agents with an $A/D > 1$ can present a genuine developmental hazard.

Table 5.1
Developmental Toxicity Endpoints

Type I	
↓Number of live births (litter size)	↓Number of live fetuses (litter size)
↑Number of still births	↑Number of resorptions
	↑Number of fetuses with malformations, structural changes
Type II	
↓Birth weights	
↓Postnatal survival	↑Number of fetuses with retarded development (skeletal, soft tissue)
↓Postnatal growth, reproductive capacity	

Source: Frankos (1985).

Since maternal toxic effects are generally non-specific and readily detectable in human populations, substances with a low A/D ratio are considered relatively less hazardous. Conversely, substances with a high A/D ratio exhibit fetotoxic effects at doses much lower than the maternal toxic dose range. These agents are considered much more hazardous developmental toxicants, because exposure limits which are adequate to protect adults from adverse health effects would be insufficient to protect the exposed fetus from developmental toxicity. For example, thalidomide has a large A/D ratio of greater than 30, while vitamin A has an A/D ratio of about two. Both substances are teratogenic, but only doses that approach the adult toxicity level will produce fetotoxic effects from vitamin A, while much smaller doses of thalidomide are powerfully teratogenic.

A variety of dose-response measures that precisely define the A/D ratio have been proposed, although none has received widespread acceptance to date (Bass *et al.*, 1985; Fabro, 1985; Johnson, 1989). A more explicit quantitative definition of the A/D relation has been proposed as the "*Relative Teratogenic Index*" (RTI), calculated as the ratio of the maternal LD_{01} for adult lethality versus the embryo TD_{05} for teratogenicity (Fabro, 1985). This measure could be included as an additional criterion in determining safety factors for teratogens, but it has not yet been widely accepted and has received criticism on methodological grounds (Johnson, 1989).

The FDA procedure for determining safety factors for teratogens specifies safety factor values between 100 and 1000 depending primarily on a "developmental specificity" matrix (Table 5.2). This procedure ascertains whether a Type I or Type II developmental toxicity endpoint has been detected either: (1) in a simple teratogenicity test, or (2) in a multigenerational (two-generational) test. If Type I effects are found in both classes of tests, then a safety factor of 1000 is assigned. If exclusively Type II effects or a combination of Type II effects and no effects are found, then a safety factor of 100 is applied. If a combination of Type I effects and the other developmental outcomes are found, then intermediate safety factor values between 100-1000 are determined using additional qualitative criteria comparable to the EPA Guidelines.

Table 5.2
Safety Factors Applied to Outcomes of Developmental Toxicity Texts

	Findings from multigeneration test		
Teratology test	Type I	Type II	no effect
Type I	Type I (1000×)	Type I ( )*	Type I ( )*
Type II	Type I ( )*	Type II (100×)	Type II (100×)
No effect	Type I ( )*	Type II (100×)	No Effect (100×)

*An appropriate safety factor will be determined after consideration of additional information on metabolism, pharmacokinetics, natural toxicity, mechanisms, etc.

Source: Frankos (1985).

The FDA procedure may be capable of producing more systematic values for safety factors than the EPA method, although it is useful mainly for well-characterized developmental toxicants that have available datasets both from simple and multigenerational developmental outcomes.

C. Mathematical Modeling Approaches for Teratogens

The alternative approach to risk estimation — the use of mathematical extrapolation models — is not currently accepted by the 1986 EPA Guidelines:

> *At present there are no mathematical models that are generally accepted for estimating developmental toxicity responses below the applied dose range. This is due to a lack of understanding of the biological mechanisms underlying developmental toxicity, intra/interspecies differences in the types of developmental events, the influence of maternal effects on the dose-response curve, and whether or not a threshold exists below which no effect would be produced by an agent.*
>
> EPA (1986c)

Although detailed mathematical models for teratogenic processes have not been adequately developed, general mathematical extrapolation procedures can be devised, such as those proposed by Gaylor (1983, 1988), Crump (1984), and Dourson (1986) for the dose-response analysis of systemic toxicants. Such mathematical procedures would be suitable for dose-response analysis of binary count data (or ratios of count data), such as litter size or malformations per litter. The Gaylor procedure is somewhat analogous to a similar low-dose extrapolation method used for carcinogenesis risk estimation, but in the present context it does not preclude the existence of a true threshold.

The main advantage of utilizing mathematical risk estimation models is that they permit a more objective calculation of low-dose risk, because case-by-case assumptions about the size of the safety factor are not required. Since such methods provide an estimate of the upper limit of risk at low doses of teratogen, they can be considered conservative in their assumptions, and therefore can produce quantitative values comparable to the safety factor approach (Gaylor, 1983). Comparing the relative magnitude of risk calculated from the NOEL-SF procedure approach versus the mathematical extrapolation procedure would also help to validate NOEL-SF risk estimates.

D. Interspecies Extrapolation

Methodological issues pertaining to interspecies extrapolation of dose-response data from test animals to humans are not specifically addressed in 1986 EPA Guidelines. Instead, it is recommended that interspecies extrapolation should be performed on a case-by-case basis, on the basis of biological similarities between test species and humans:

> *The determination of the size of safety factors for interspecies conversion is not a mathematical problem, but rather one involving biological interpretation.*
>
> IRLG (1986)

Evidence for the validity of animal test results in predicting teratogenic effects in humans is not extensive, but there is a good general concordance between animal and human embryotoxicity for most substances:

> *In these comparisons (animal-human), there was almost always concordance of effects between humans and at least one animal species tested. Thus, there is some limited basis for estimating the risk of exposure to human development based on data from animal studies.*
>
> EPA (1986c)

In general, developmental toxicity responses in animals and humans appear to be similar both in a qualitative and quantitative sense, so that human embryos are not significantly more susceptible to teratogenic effects than animal embryos (Johnson, 1989). Thus the magnitude of the safety factor required to account for interspecies differences in sensitivity to teratogenic agents should not be greater than for systemic toxicants, although the unpredictability of teratogenic effects between species will occasionally produce a significantly higher risk in humans (e.g. thalidomide) than predicted from animal studies.

E. Special Problems of Safety Factor Determination for Teratogens

Risk assessment methods for teratogenic hazards include several special features not generally considered in conventional risk assessments. Simple measures of the level and duration of maternal exposure to a teratogenic substance do not necessarily provide a reliable predictor of the magnitude of the teratogenic response at target sites within the developing embryo, because the dose-response relationship is closely related to the stage of development. Thus a single brief exposure may be sufficient to produce adverse developmental effects during a critical period of organogenesis, and doses that produce chronic toxicity in the mother are often not required to produce a teratogenic outcome (WHO, 1984).

In addition, the specific high-risk group for exposure to teratogens includes not only women with visible pregnancies, but all fertile women of reproductive age, since an undetected pregnancy during the period of early embryogenesis is a frequent occurrence. Indeed, the period of greatest susceptibility to teratogenic effects occurs in the first 4 weeks of pregnancy, when many women are not yet aware that they have conceived. Thus precautionary measures such as removal from high-exposure locations for women who are visibly pregnant will not protect the embryos of fertile women in the earliest stages of pregnancy. Consequently, selective criteria for exposure of fertile women to potential teratogenic substances may have to be considered, especially for occupational standard-setting. Such special considerations pose serious problems with respect to the work equity and sex discrimination provisions of labor legislation.

F. Summary of Safety Factor Determination for Teratogens

For risk estimation, the primary method for determining permissible exposure limits for teratogenic effects should be the NOEL-SF approach, in accordance with the 1986 EPA Guidelines. However, a maximum exposure limit could also be presented, based on an alternative mathematical extrapolation procedure, for those endpoints amenable to such an approach, i.e., for endpoints with count and ratio data. Selection of either the NOEL-SF or mathematical extrapolation method should be left to scientific judgement, according to the biostatistical properties of the dose-response data and the relevant developmental endpoints demonstrated in the animal experiments.

Estimated NOELs and mathematical extrapolation curves tend to exhibit large error bounds, due to the small sample size for determining individual points in the dose-response curve, and the large variability in the teratogenic responses of the test animals. Therefore, teratogenic risk should be estimated at the 95% upper confidence limit both for the NOEL-SF method and the mathematical extrapolation method. A best-estimate for the NOEL or a best-fit for the extrapolation curve might also be presented for comparative purposes, but these values should not be employed for the purposes of deriving exposure limits.

The greatest source of uncertainty in the estimation of exposure limits for teratogens using the NOEL-SF approach is the selection of an appropriate value for the safety factor. The risk assessment should specify the magnitude of the safety factor employed, and should discuss the assumptions used and accuracy of estimates in determining this value. Normally a minimum value of 100 is selected, but larger safety factors may be chosen according to the special criteria previously discussed.

References

ACGIH (1986). American Conference of Government Industrial Hygienists. *Identification and Classification of Carcinogens: Procedures of the Chemical Substances Threshold Limit Value Committee*. ACGIH. Am J Public Health 76: 1232-1235.

Aldridge WN (1986). *The Biological Basis and Measurement of Thresholds*. Ann Rev Pharmacol Toxicol 26: 39-58.

Andersen ME, Clewell HJ, Gargas ML, Smith FA, Reitz RH (1987). *Physiologically Based Pharmacokinetics and the Risk Assessment Process for Methylene Chloride*. Toxicol Appl Pharmacol 87: 185-205.

Anderson EL (1985). *Quantitative Approaches in Use in the United States to Assess Cancer Risk*, in: Vouk VB et al, Eds., Methods for Estimating Risk for Chemical Injury: Human and Non-Human Biota and Ecosystems (SCOPE 26). John Wiley and Sons, Toronto, pp 405-436.

Ashford NA, Ryan CW, Caldart CC (1983). *Law and Science Policy in Federal Regulation of Formaldehyde*. Science 222: 894-900.

Barnes DG and Dourson ML (1988). *Reference Dose (RfD): Description and Use in Health Risk Assessments* Regul Toxicol Pharmacol 8: 471-486.

Bass R, Neubert D, Stotzer H, Bochert G (1985). *Quantitative Dose-Response Models in Prenatal Toxicology*, in Vouk VB *et al.*, Eds., Methods for Estimating Risk of Chemical Injury: Human and Non-Human Biota and Ecosystems (SCOPE 26). John Wiley and Sons, Toronto, pp 437-456.

Bigwood EJ (1973). *The Acceptable Daily Intake of Food Additives*. CRC Crit Rev Toxicol (June): 41-93.

Calabrese EJ (1978a). *Methodological Approaches to Deriving Environmental and Occupational Health Standards*. John Wiley and Sons, New York.

Calabrese EJ (1987b). *Pollutants and High Risk Groups*. John Wiley and Sons, New York.

Calabrese EJ (1983). *Principles of Animal Extrapolation*. John Wiley and Sons, New York.

Calabrese EJ (1985). *Uncertainty Factors and Interindividual Variation*. Regul Toxicol Pharmacol 5: 190-196.

Calabrese EJ (1987). *Animal Extrapolation: A Look Inside the Toxicologist's Black Box*. Environ Sci Technol 21:618-623.

Cotruvo JA (1988). *Drinking Water Standards and Risk Assessment*. Regul Toxicol Pharmacol 8: 288-299.

Crump KS (1984). *A New Method For Determining Allowable Daily Intakes*. Fundam Appl Toxicol 4: 854-871.

Crump KS (1986). *Letter*. Fundam Appl Toxicol 6: 183-184.

Dourson ML (1986). *New Approaches in the Derivation of Acceptable Daily Intake (ADI)*. Comments Toxicology 1:35-48.

Dourson ML, Stara JF (1983). *Regulatory History and Experimental Support of Uncertainty (Safety) Factors*. Regul Toxicol Pharmacol 3: 224-238.

Dourson ML, Hertzberg RC, Hartung R, Blackburn K (1985). *Novel Method for the Estimation of the Acceptable Daily Intake*. Toxicol Ind Health 1:23-41.

Dourson ML, Hertzberg RC, Stara JF (1986). *Letter*. Fundam Appl Toxicol 6: 182-183.

EPA (1980). U.S. Environmental Protection Agency. *Guidelines and Methodology Used in the Preparation of Health Effects Assessment Chapters of the Consent Decree Water Quality Criteria*. Federal Register 45: 79347-79357.

EPA (1985). U.S. Environmental Protection Agency. *National Primary Drinking Water Regulations; Synthetic Organic Chemicals, Inorganic Chemicals, and Microorganisms*. Federal Register 50: 46936-47025.

EPA (1986a). U.S. Environmental Protection Agency. *Guidelines for the Health Risk Assessment of Chemical Mixtures*. Federal Register 51: 34014-34025.

EPA (1986b). U.S. Environmental Protection Agency. *Guidelines for Carcinogen Risk Assessment*. Federal Register 51: 33992-34003.

EPA (1986c). U.S. Environmental Protection Agency. *Guidelines for the Health Assessment of Suspect Developmental Toxicants.* Federal Register 51: 34028-34040.

EPA (1987). *A Cancer Risk Specified Dose Estimate for 2,3,7,8-TCDD : Review Draft.* Environmental Protection Agency, Washington DC.

Fabro S (1985). *On Predicting Environmentally-Induced Human Reproductive Hazards: An Overview and Historical Perspective.* Fundam Appl Toxicol 5: 609-614.

Finkel AM (1988). *Dioxin: Are We Safer Now Than We Were Before?* Risk Analysis 8: 161-165.

Frankos VH (1985). *FDA Perspectives on the Use of Teratology Data for Human Risk Assessment.* Fundam Appl Toxicol 5: 615-625.

Gaylor DW (1983). *The Use of Safety Factors for Controlling Health Risks.* J Toxicol Environ Health 11: 329-336.

Gaylor DW (1988). *Applicability of Cancer Risk Assessment Techniques to Other Effects.* Toxicol Ind Health 4: 453-459.

Gaylor DW, Kodell RL (1980). *Linear Interpolation Algorithm for Low Dose Risk Assessment of Toxic Substances.* J Environ Pathol Toxicol 4: 305-312.

Gough M (1988). *Science Policy Choices and the Estimation of Cancer Risk Associated with Exposure to TCDD.* Risk Analysis 8: 337-342.

Hartung R, Durkin PR (1986). *Ranking the Severity of Toxic Effects: Potential Applications to Risk Assessment.* Comments Toxicology 1:49-63.

Hattis D (1987). *The Value of Molecular Epidemiology in Quantitative Health Risk Assessment,* in: Draggan S et al., Eds., Environmental Impacts on Human Health. Preager, New York.

Hattis D, Erdreich L, Ballew M (1987). *Human Variability in Susceptibility to Toxic Chemicals — A Preliminary Analysis of Pharmacokinetic Data from Normal Volunteers.* Risk Analysis 7: 415-426.

Hoel DG, Kaplan NL, Anderson MW (1984). *Implication of Non-linear Kinetics on Risk Estimation in Carcinogenesis.* Science 219: 1032-1036.

Hutt PB (1985). *Use of Quantitative Risk Assessment in Regulatory Decisionmaking under Federal Health and Safety Statutes,* in: DG Hoel *et al.*, Eds., Risk Quantitation and Regulatory Policy (Banbury Report 19). Cold Spring Harbor Laboratory, New York.

IRLG (1979). Interagency Regulatory Liaison Group, Work Group on Risk Assessment. *Scientific Bases for the Identification of Potential Carcinogens and Estimation of Risks.* Federal Register 44: 39858-39879. [également publié dans: Annual Review of Public Health 1: 345-393 (1980)].

IRLG (1986). Interagency Regulatory Liaison Group (Kimmel CA, Kimmel GL, and Frankos VH, Eds.). *Interagency Regulatory Liaison Group Workshop on Reproductive Toxicity Risk Assessment.* Environ Health Perspect 66: 193-221.

Johnson EM (1989). *A Case Study of Developmental Toxicity Risk Estimation Based on Animal Data: The Drug Bendectin,* in: DJ Paustenbach, Ed., The Risk Assessment of Environmental and Human Health Hazards: A Textbook of Case Studies. John Wiley and Sons, New York.

Kimmel CA, Gaylor DW (1988). *Issues in Qualitative and Quantitative Risk Analysis for Developmental Toxicology.* Risk Analysis 8:15-20.

Kociba RJ, Keyes DG, Beyer JE, et coll. (1978). *Results of a Two-year Chronic Toxicity and Oncogenicity Study of 2,3,7,8-TCDD in rats.* Toxicol Appl Pharmacol 46: 279-303.

Krewski D, Brown C, Murdoch D (1984). *Determining "Safe" Levels of Exposure: Safety Factors or Mathematical Models?* Fundam Appl Toxicol 4: S383-S394.

Kroes R, Feron VJ (1984). *General Toxicity Testing: Sense and Nonsense, Science and Policy.* Fundam Appl Toxicol 4: S298-S308.

Lehman AJ, Fitzhugh OG (1954). *100-fold Margin of Safety.* Assoc Food Drug Off U.S. Q Bull 18: 33-35.

Lu FC (1983). *Toxicological Evaluations of Carcinogens and Non-Carcinogens: Pros and Cons of Different Approaches.* Regul Toxicol Pharmacol 3:121-132.

Lu FC (1985). *Safety Assessments of Chemicals with Thresholded Effects.* Regul Toxicol Pharmacol 5: 460-464.

Lu FC (1988). *Acceptable Daily Intake: Inception, Evolution, and Application.* Regul Toxicol Pharmacol 8: 45-60.

Mantel N (1975). *The Assessment of Carcinogenicity.* Cancer Research 35:865-876.

Marshall E (1982). *EPA's High Risk Carcinogen Policy.* Science 218: 975-982.

Munro IC, Krewski DR (1981). *Risk Assessment and Regulatory Decision-Making.* Fd Cosmet Toxicol 19: 549-560.

NAS (1977). *Drinking Water and Health.* National Research Council, U.S. National Academy of Sciences. Washington DC.

Nelson N (1984). *Use of Toxicity Test Data in the Estimation of Risks to Human Health,* in: Rodricks JV and Tardiff RG, Eds., Assessment and Management of Chemical Risks (ACS Symposium Series 239). American Chemical Society, Washington.

OECD (1981). Organization for Economic Co-Operation and Development. *OECD Guidelines for Testing of Chemicals #414 -Teratogenicity.*

OSTP (1986). Office of Science and Technology Policy, U.S. Interagency Staff Group on Carcinogens (Hart RW, Chairman). *Chemical Carcinogenesis: A Review of the Science and its Associated Principles*. Environ Health Perspect 67: 201-282.

Perera FP (1988). *EPA Cancer Risk Assessment (letter)*. Science 239: 1227.

Perera FP, Petito C (1982). *Formaldehyde: A Question of Cancer Policy*. Science 216: 1285-1291.

Preuss PW, Erlich AM (1987). *Environmental Protection Agency's Risk Assessment Guidelines. Health of Non-Criteria Pollutants*. J Air Pollution Control Assoc 37: 783-797.

Rodricks JV, Tardiff RG (1984). *Conceptual Basis for Risk Assessment*, in: Rodricks JV and Tardiff RG, Eds., Assessment and Management of Chemical Risks (ACS Symposium Series 239). American Chemical Society, Washington.

Rushefsky ME (1985). *Assuming the Conclusions: Risk Assessment in the Development of Cancer Policy*. Politics and the Life Sciences 4: 3144.

Sharratt M (1976). *Uncertainties Associated with the Evaluation of the Health Hazards of Environmental Chemicals from Toxicological Data*, in: Hunter WJ and Smeets JG, Eds., The Evaluation of Toxicological Data for the Protection of the Environment, Pergamon Press, Toronto.

Shu HP, Paustenbach DJ, Murray FJ (1987). *A Critical Evaluation of the Use of Mutagenesis, Carcinogenesis and Tumor Promotion Data in a Cancer Risk Assessment of 2,3,7,8-Tetrachlorodibenzo-p-Dioxin*. Regul Toxicol Pharmacol 7: 57-88.

Somers E (1981). *Environmental Risk Management in Canada*. Regul Toxicol Pharmacol 3: 75-81.

Somers E (1984). *Risk Estimation for Environmental Chemicals as a Basis for Decision-making*. Regul Toxicol Pharmacol 4: 99-106.

Stara JF, Bruins RJF, Dourson ML, Erdreich LS, Hertzberg RC, Durkin PR, Pepelko WE (1987). *Risk Assessment is a Developing Science: Approaches to Improve Evaluation of Single Chemicals and Chemical Mixtures*, in: VB Vouk et al, Eds., Methods for Assessing the Effects of Mixtures of Chemicals (SCOPE 30). John Wiley and Sons, Toronto.

Vettorazzi G (1976). *Safety Factors and their Application in the Toxological Evaluation*, in: WJ Hunter and JG Smeets, Eds., The Evaluation of Toxicological Data for the Protection of Public Health. Pergamon Press, Oxford.

Vocci F, Farber T (1988). *Extrapolation of Animal Toxicity Data to Man*. Regul Toxicol Pharmacol 8: 389-398.

Watanabe PG, Schumann AM, Reitz RH (1988). *Toxicokinetics in the Evaluation of Toxicity Data*. Regul Toxicol Pharmacol 8: 408-413.

Weil CS (1972). *Statistics versus Safety Factors and Scientific Judgment in the Evaluation of Safety Factors for Man*. Toxicol Appl Pharmacol 21: 454-463.

Weil CS, McCollister DD (1963). *Relationship between short- and long-term Feeding Studies in Designing an Effective Toxicity Test*. Agric Food Chem 11: 486-491.

Weisberger JH and Williams GM (1981). *Carcinogen Testing: Current Problems and New Approaches*. Science 214: 401-407.

WHO (1984). World Health Organization. *Principles for Evaluating Health Risks to Progeny Associated with Exposure to Chemicals During Pregnancy*. Environmental Health Criteria 30. World Health Organization, Geneva.

WHO (1987). World Health Organization. *Principles for the Safety Assessment of Food Additives and Contaminants in Food*. Environmental Health Criteria 70. World Health Organization, Geneva.

Facteurs de sécurité biologiques et évaluation du risque toxicologique

Santé et Bien-être social Canada
Health and Welfare Canada
Canada

90-DHM-154

Facteurs de sécurité biologiques et évaluation du risque toxicologique

Rapport préparé pour la

Direction de l'hygiène du milieu
Direction générale de la protection de la santé
Santé et Bien-être social Canada

par
R. Stephen McColl, Ph.D.
du Département des études sur la santé
et de
l'*Institute for Risk Research*
Université de Waterloo

Also available in English under the title:
"Biological Safety Factors in
Toxicological Risk Assessment"

On peut obtenir des copies de ce rapport auprès de la:
Direction générale des communications
Ministère de la Santé nationale et du Bien-être social
Immeuble Brooke-Claxton
5e étage
Ottawa (Ontario)
K1A 0K9

© Ministre des Approvisionnements et Services Canada
Cat. 49-49/1990F
ISBN 0-662-95920-5

Ce rapport est le résultat d'un travail de recherche commandé par la Direction de l'hygiène du milieu, Direction générale de la protection de la santé, Santé et Bien-être social Canada.

L'auteur de ce rapport, R. Stephen McColl, Ph.D., est membre de l'*Institute for Risk Research* et membre du corps enseignant au Département des études sur la santé de l'Université de Waterloo.

Les docteurs D. Krewski et V. Armstrong, de la Direction de l'hygiène du milieu, ministère de la Santé nationale et du Bien-être social, ont révisé ce rapport et y ont apporté des corrections.

Les avis et les recommandations exprimés dans le présent document reflètent l'opinion de l'auteur et pas nécessairement celui du ministère de la Santé nationale et du Bien-être social. Les résultats, les modèles, les calculs et les recommandations présentés dans cette étude figurent à titre explicatif et ne doivent pas servir à l'élaboration de règlements.

Mars 1989

Table des matières

	Page
1. Notions de base pour la détermination des facteurs de sécurité	**7**
A. La notion de seuil	7
B. Classification des effets toxiques	8
2. Dose journalière admissible et facteurs de sécurité	**10**
A. Élaboration de la notion de facteur de sécurité	10
B. Méthodologie relative au facteur de sécurité DSEO	12
1. Définition de la DJA	12
2. Définition de la DSEO et d'autres estimateurs de seuil	13
C. Description des facteurs de sécurité	15
1. Définition d'un facteur de sécurité	15
2. Choix des facteurs de sécurité	15
D. Validation des facteurs de sécurité	17
1. Facteur global de sécurité	17
2. Facteurs spécifiques de sécurité	17
a) Variation intra-espèce	17
b) Extrapolation inter-espèces	18
c) Correction subchronique/chronique	18
d) Correction DMEO/DSEO	19
3. Autres conceptions des méthodes FS	**20**
A. Critiques de la méthode DSEO-FS	20
B. Modèles mathématiques pour le facteur de sécurité	21
1. Modèles de distribution normale (probit)	21
2. Méthode d'extrapolation linéaire de Gaylor	22
3. Méthode du facteur de sécurité de Dourson	23
4. Méthode de la dose-repère de Crump	23
5. Résumé et critique	25
C. Gravité des effets	25
1. Échelles de gravité	25
2. Estimation de DSEO au moyen de diagrammes dose-gravité	26

	Page
4. Méthodes FS appliquées aux produits cancérigènes	**28**
A. Généralités	28
B. Facteurs traditionnels de sécurité pour produits cancérigènes	28
C. Méthodes de classification selon l'importance des preuves	29
D. Système de classification mécanistique	31
E. Méthodes mécanistiques d'évaluation du risque appliquées au cas de la dioxine	33
1. Méthodologie de l'EPA	33
2. Méthodologie de l'OME	34
F. Modélisation PBP	35
5. Méthodes FS appliquées aux tératogènes	**38**
A. Généralités	38
B. Méthode DSEO-FS appliquée aux tératogènes	39
C. Méthodes mathématiques de modélisation appliquées aux tératogènes	41
D. Extrapolation inter-espèces	41
E. Problèmes spéciaux reliés à la détermination des FS pour les tératogènes	41
F. Résumé	42
Bibliographie	**43**

1. Notions de base pour la détermination des facteurs de sécurité

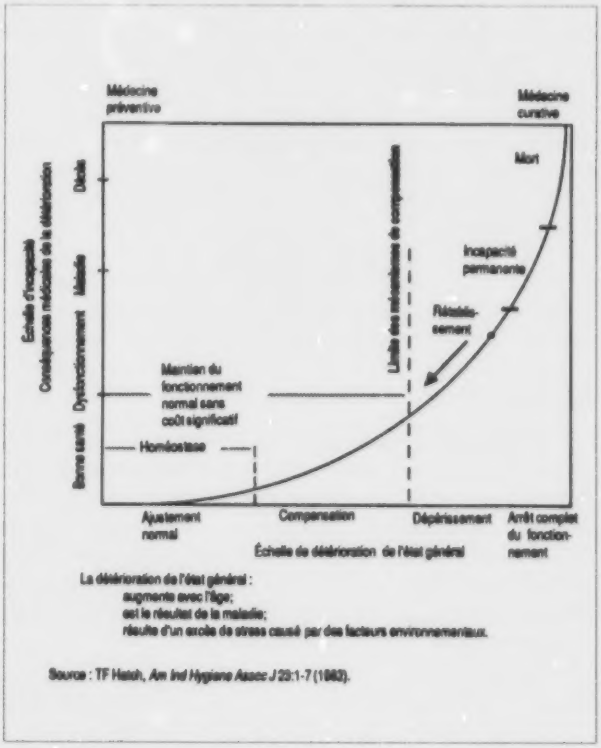
A. La notion de seuil

Les méthodes traditionnelles d'évaluation de la toxicité causée par des facteurs environnementaux mettaient surtout l'accent sur la détermination des niveaux «sûrs» d'exposition à des substances toxiques. Dans le présent contexte, la «sécurité» se définit comme l'absence d'effets toxiques pour la population humaine exposée. Toutefois, en théorie et en pratique, les études toxicologiques ne peuvent fournir que des estimations des concentrations minimales de risque, telles que définies de façon opérationnelle par des épreuves effectuées sur des animaux ou par des études épidémiologiques; les études toxicologiques ne sauraient prévoir des niveaux correspondant à la sécurité absolue (Sharratt, 1976; Kroes et Feron, 1984; Rodricks et Tardiff, 1984).

L'estimation des conditions sûres d'exposition à la plupart des toxiques est fondée sur la *notion de seuil*, selon laquelle le niveau posologique détermine si l'exposition entraînera ou non un effet pathologique. D'après la théorie toxicologique traditionnelle, un organisme peut tolérer de faibles doses de toxiques grâce à la présence de métabolismes de détoxication, de mécanismes d'homéostase physiologique, de l'adaptabilité des cellules et de systèmes de réparation. En-deçà d'un certain niveau minimal, ces mécanismes de compensation peuvent contrer les effets d'un toxique; par conséquent, on peut définir le seuil toxicologique comme la dose en-deçà de laquelle l'organisme ne subit aucune modification pathologique (Aldridge, 1986). Lorsque le toxique est présent à des doses plus élevées, le système de défense de l'organisme finit par ne plus suffire à la tâche, ce qui entraîne la perte de l'homéostase physiologique et la détérioration des fonctions organiques (Figure 1.1). Finalement, la détérioration physiologique se manifeste sous la forme d'un ou de plusieurs processus pathologiques qui causent la maladie dans l'organisme exposé. Le niveau de ce seuil détermine le *point critique*.

Dans la pratique, le processus de détermination des seuils de toxicité se révèle en fait plus compliqué: en effet, à mesure que la dose augmente, les effets toxiques se manifestent tour à tour au niveau de plusieurs organes cibles, puisque les divers mécanismes de compensation ou de défense se trouvent neutralisés à des doses différentes (Hattis, 1987). Par conséquent, il faut définir plusieurs seuils différents en fonction de la complexité de l'organisme humain. Pour chaque point critique, on doit déterminer un *seuil individuel* défini comme le niveau d'exposition qui n'entraîne pas d'effet toxique chez une personne donnée. De plus, on doit définir un *seuil de population* en fonction de la réaction cumulative de toutes les personnes à l'exposition d'un toxique, en tenant compte des variations naturelles de sensibilité présentes dans la population.

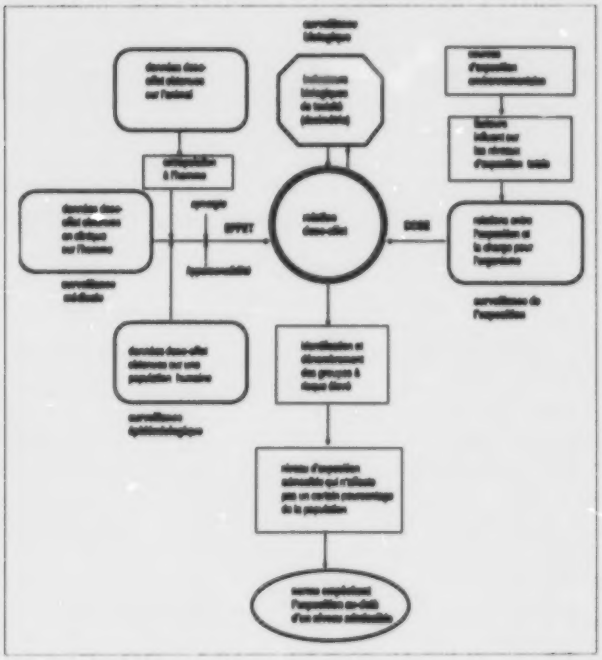
Figure 1.1
Courbe de détérioration de l'état général-incapacité



Le manque de données relatives à la toxicité d'un grand nombre de substances dangereuses pour l'homme crée souvent de l'incertitude quant à la répartition des seuils de toxicité dans une population, particulièrement pour les personnes les plus

sensibles. De plus, les expositions multiples à des mélanges de toxiques présents dans l'environnement peuvent causer des effets synergiques ou potentialisés imprévisibles, augmentant ainsi de façon significative la toxicité des produits chimiques (Figure 1.2). Enfin, les animaux de laboratoire servent souvent à la détermination de la toxicité des substances: par conséquent, toute valeur de seuil tirée des expériences effectuées sur les animaux doit être convertie en son équivalent correspondant à l'exposition de l'homme au produit étudié, à l'aide de méthodes d'extrapolation inter-espèces dont l'exactitude dépend de plusieurs suppositions et incertitudes.

Figure 1.2
Modèle pour l'élaboration des normes d'exposition



B. Classification des effets toxiques

Les dangers environnementaux n'affectent pas toujours la santé selon le même type de mécanisme pathologique de base. En conséquence, le choix d'une méthode donnée d'évaluation du risque dépendra de la classification du mécanisme d'action de la toxicité. Hattis (1987) a décrit un schéma de classification pratique, selon lequel les effets toxiques sont répartis en deux grandes classes d'après la causalité des mécanismes de maladie et la réversibilité des effets (Tableau 1.1).

Tableau 1.1
Types de dangers pour la santé qui exigent des méthodes d'évaluation du risque fondamentalement différentes

Toxicité déterministe «Traditionnelle»

Procède en neutralisant les processus de compensation de l'organisme; au-dessous d'un certain seuil, chez des personnes qui ne sont pas encore au-delà des limites du fonctionnement normal en absence d'exposition, la réaction est réversible.

- Toxicité aiguë traditionnelle
  L'effet toxique est complètement réversible ou entraîne des lésions à long terme dans les trois jours de l'exposition, par exemple, des lésions pulmonaires à la suite d'inhalation de chlore.
- Toxicité chronique traditionnelle
  La toxicité entraîne habituellement des lésions permanentes au cours d'une période pouvant varier de plusieurs jours à plusieurs mois, en raison: a) de l'accumulation d'un produit toxique (par exemple, le méthyle mercure), ou b) de l'accumulation des effets de réactions peu réversibles à des produits toxiques.

Toxicité non déterministe

Ses effets résultent de mécanismes insidieux qui sont irréversibles ou peu réversibles à des doses faibles ou au cours des premières phases du déclenchement.

- Effets au niveau de la biologie moléculaire (Processus stochastique)
  – ont lieu lorsqu'il y a un ou quelques changements irréversibles au niveau de l'information codée dans l'ADN, par exemple la mutagénèse, la plupart des cas de carcinogénèse et certains cas de tératogénèse.
- Effets cumulatifs chroniques
  – ont lieu par suite de l'accumulation chronique d'un grand nombre de lésions à petite échelle, par exemple les malformations congénitales, l'emphysème, la surdité causée par le bruit, l'athéromatose et probablement l'hypertension.

Adaptation de Hattis, 1987.

L'évaluation de la toxicité «traditionnelle» est fondée sur un mécanisme *déterministe* de la toxicité. Les processus déterministes agissent en neutralisant les mécanismes de compensation de l'organisme exposé à un toxique à une dose suffisamment forte: le seuil de toxicité est alors dépassé et les effets pathologiques se manifestent. Ces effets sont toutefois réversibles jusqu'au cours des phases tardives du processus pathologique (Figure 1.1). Il est alors question de *toxicité systémique*, qui comprend à la fois la toxicité aiguë et la toxicité chronique traditionnelles, dans le cas des effets sur la santé qui ne sont pas associés au cancer ou à la mutation d'un gène (Barnes et Dourson, 1988). Les mécanismes pathologiques inhérents à la toxicité systémique peuvent être résumés ainsi:

Lorsqu'une fonction biologique qui dépend de l'action concertée d'un grand nombre de macromolécules identiques comme les récepteurs ou les enzymes est perturbée, l'existence d'un seuil se révèle alors probable. Dans ce cas, l'action d'un certain nombre de molécules de récepteur ou d'enzyme peut être inhibée ou modifiée sans que la fonction biologique ne soit affectée de façon notable. Cependant, lorsque la fonction biologique dépend de l'action d'une molécule comme l'ADN ou de l'une des molécules (enzymes) qui y sont associées selon un rapport de un à un, l'existence d'un seuil ne semble plus vraisemblable. De ce point de vue, l'exposition à la plupart des toxiques à faible concentration au cours d'une vie sera sans danger. À l'opposé, on peut présumer de la nocivité des produits chimiques qui entraînent des modifications génétiques. De tels énoncés ne tiennent pas compte des questions de risque mettant en jeu la pente des courbes dose-effet ... ou des différences de susceptibilité qui existent d'une personne à une autre. Des décisions d'ordre pratique doivent être prises même si le risque n'est pas nul.

Aldridge (1986)

En supposant l'existence d'un mécanisme de seuil, on a fondé la méthode traditionnelle d'évaluation des produits de toxicité générale sur la notion de *dose journalière admissible* (DJA). Il s'agit d'un paramètre toxicologique défini comme la quantité de substance que l'organisme peut absorber quotidiennement au cours d'une période de temps prolongée sans subir d'effets délétères. Le calcul de la DJA comprend habituellement un *facteur de sécurité* (FS), qui tient compte de l'incertitude inhérente à l'estimation du seuil de toxicité. Toutefois, l'élaboration de méthodes uniformisées d'évaluation du risque est récente; cette uniformisation est nécessaire à la cohérence de l'interprétation des expériences toxicologiques, surtout dans le cas des effets généraux chroniques. Auparavant, on évaluait les effets des toxiques systémiques sur la santé à l'aide d'expériences toxicologiques généralement acceptées mais pas nécessairement bien validées. Une grande partie de l'interprétation des données toxicologiques a ainsi été laissée à «l'expertise» des toxicologues, ce qui a entraîné des variations méthodologiques parmi les organismes qui élaboraient des règlements.

La deuxième classe d'effets toxiques provient de processus *non déterministes* qui n'adoptent pas le comportement dose-effet caractéristique aux mécanismes de toxicité traditionnelle (Hattis, 1987). Ces processus non déterministes ne présentent pas de seuil évident de toxicité: ainsi, l'exposition prolongée à de faibles doses de toxique peut entraîner des effets toxiques graves qui ne sont pas nécessairement proportionnels aux paramètres d'exposition. De plus, ces processus sont irréversibles ou peu réversibles au cours des phases précoces de l'atteinte.

Cette classe comprend les maladies comme le cancer et les mutations géniques résultant de processus *stochastiques*. Dans ce cas, la probabilité de l'occurrence d'une maladie est déterminée par la probabilité de l'existence d'une atteinte au niveau d'une molécule cible dans une cellule, plus particulièrement une modification de l'information génétique contenue dans l'ADN. Les processus stochastiques agissent selon un schéma probabiliste d'effets tout-ou-rien, de sorte que la dose ne détermine que la probabilité d'occurrence d'une maladie et non la gravité de la maladie elle-même. En principle, les processus stochastiques ne comportent pas de seuil et ne sont par conséquent pas habituellement traitables selon la méthode de la DJA.

Les autres effets chroniques cumulatifs qui ne sont pas nécessairement stochastiques peuvent aussi se manifester de façon non déterministe à la suite de l'accumulation au cours des années d'un grand nombre de lésions ou d'épisodes traumatisants à petite échelle, comme l'emphysème pulmonaire, la surdité causée par le bruit, ou les problèmes neuronaux causés par l'exposition au plomb chez les jeunes enfants. De plus, un grand nombre de perturbations de la régulation cellulaire, dont la plupart des cas de tératogénèse et certains cas de carcinogénèse (par exemple, les produits promoteurs de tumorisation), peuvent survenir à cause d'effets toxiques qui dépendent plus du schéma d'exposition et des moments précis auxquels elle a eu lieu, que de la dose elle-même. Cependant, en tenant compte du mécanisme d'action spécial et de la cinétique d'exposition inhabituelle, on peut quand même tenter d'évaluer le risque selon la méthode DJA pour ces effets; on ne doit toutefois pas oublier que la validité et l'exactitude de ces évaluations sont moins bonnes que dans le cas des processus déterministes traditionnels.

2. Dose journalière admissible et facteurs de sécurité

A. Élaboration de la notion de facteur de sécurité

La méthodologie de la détermination de la *dose journalière admissible* (DJA) est fondée principalement sur des expériences dose-effet menées sur des animaux de laboratoire, ou plus rarement sur des expositions bien documentées chez l'homme. Pour chaque type d'effet résultant d'une exposition prolongée à un produit chimique spécifique, la méthodologie DJA étudie la relation dose-effet dans le but d'établir la *dose qui n'entraîne pas d'effet observé* (DSEO) pour un effet biologique spécifique (Lu, 1983, 1985, 1988; Dourson et Stara, 1983). En raison des incertitudes inhérentes à l'extrapolation inter-espèces de la DSEO de l'animal à l'homme, et en raison de la variabilité de la sensibilité humaine aux effets des toxiques, il est impératif d'introduire un facteur de sécurité qui abaisse la valeur de la DSEO à un niveau plus sûr pour la majorité de la population humaine.

La méthode traditionnelle de calcul de la DJA pour un toxique consiste à diviser la valeur de la DSEO par un facteur de sécurité (FS) préalablement défini:

$$\text{DJA} = \text{DSEO}/\text{FS}$$

Par convention, le facteur de sécurité est le plus souvent fixé d'abord à 100, ce qui est généralement considéré suffisant pour tenir compte des incertitudes de l'extrapolation inter-espèces de l'animal à l'homme et pour tenir compte des variations individuelles de sensibilité aux produits chimiques chez l'homme (Lu, 1983). Toutefois, plusieurs autres corrections de la valeur du facteur global de sécurité peuvent être introduites en fonction de la nature et de la qualité des données dose-effet, si des renseignements toxicologiques complémentaires sont disponibles.

La notion de dose journalière admissible (DJA) a d'abord été proposée par le Joint FAO/WHO Expert Committee on Food Additives (JECFA) (Comité conjoint de l'OAA/OMS pour les additifs alimentaires) en 1961 (Lu, 1988). En 1962, les membres du Joint FAO/WHO Meeting of Experts on Pesticide Residues (JMPR) (Congrès conjoint OAA/OMS d'experts au sujet des résidus de pesticides) ont adopté le même terme et ils ont aussi décrit les méthodes de détermination de la valeur de la DJA et leur application pratique. Ces deux organismes internationaux d'experts en évaluation toxicologique des additifs alimentaires et des résidus de pesticides ont continué d'utiliser la méthodologie DJA, que les organismes chargés de l'élaboration des règlements dans un grand nombre de pays ont aussi adoptée pour la formulation des normes nationales relatives à la sécurité alimentaire (OMS, 1987).

En 1954, avant même l'adoption officielle de la notion de DJA, l'idée d'un facteur de sécurité d'une valeur de 100 avait été avancée comme «marge de sécurité appropriée» pour la dérivation de normes d'exposition aux additifs alimentaires et aux contaminants par la Food and Drug Administration des États-Unis (U.S. FDA). La marge de sécurité visait à assurer que les substances utilisées dans les aliments ne produiraient aucun effet délétère sur la santé de l'animal qui en consommait des doses au moins 100 fois supérieures à celles que l'homme était susceptible d'absorber (Calabrese, 1978a). La notion initiale de facteur de sécurité était plus intuitive que scientifique, et il était expressément mentionné que:

> *... la notion de marge de sécurité constitue une façon raisonnable d'aborder le sujet, mais son acceptation ne devrait pas amener les chercheurs et/ou le public à croire à tort que son existence a un fondement expérimental ou théorique.*
>
> J.M. Barnes et F.A. Denz (1954), dans Calabrese (1978a)

La raison d'être initiale de l'utilisation du facteur de sécurité de 100 fois par la FDA était de compenser la variabilité intra-espèce et la variabilité inter-espèces quant à la toxicité d'un produit chimique, de tenir compte des sous-groupes de sensibilité dans la population humaine, et de tenir compte de la synergie possible entre différents additifs alimentaires et contaminants. Le choix d'une valeur de 100 était fondé sur les résultats d'études toxicologiques menées à la fin des années quarante selon lesquels la sensibilité de l'homme est 10 fois supérieure à celle de l'animal de laboratoire, et selon lesquels il existe une différence d'un facteur 10 de la susceptibilité des personnes dans une population (Hutt, 1985). Globalement, la variation attendue était d'un facteur de 100 entre l'animal de

laboratoire et la personne la plus sensible. Par la suite, le facteur de sécurité de 100 a été adopté par l'OMS en 1958, et a été recommandé en 1970 comme notion de base pour la réglementation des additifs alimentaires aux États-Unis par le *Food Protection Committee of the National Research Council, National Academy of Sciences* (Comité pour la protection alimentaire du Conseil national de recherche, Académie nationale des sciences) (Lu, 1983). Plus tard, la FDA a élargi la notion de facteur de sécurité en recommandant l'inclusion d'un facteur supplémentaire de 10 dans les cas où seuls les résultats d'expériences de toxicité à court terme étaient disponibles. Le facteur global de sécurité était alors de 1000 pour les produits chimiques pour lesquels des expériences subchroniques avaient été menées sur deux espèces animales, et de 2000 pour les produits chimiques pour lesquels des expériences subchroniques avaient été menées sur une seule espèce animale (Dourson et Stara, 1983).

Aux États-Unis, la notion de DJA élaborée par la FDA a par la suite été adoptée pour la définition de normes réglementaires dans d'autres secteurs de l'hygiène du milieu. Beaucoup de mises au point consécutives de l'utilisation de facteurs de sécurité pour la détermination de la DJA proviennent originalement de l'élaboration de normes de qualité de l'eau ambiante quant aux polluants présents dans l'eau potable. En 1977, le *Safe Drinking Water Committee* (Comité pour de l'eau potable sans danger) de la *National Academy of Sciences* (NAS) des États-Unis présentait un rapport d'un grand intérêt intitulé: *Drinking Water and Health* (Eau potable et santé) (NAS, 1977), qui comportait des révisions non négligeables des méthodes de calcul des DJA relatives aux normes américaines pour l'eau potable.

Les recommandations visant l'évaluation du risque dans *Drinking Water and Health* différaient des méthodes préconisées par la FDA sous deux aspects. En premier lieu, la NAS proposait l'expression d'une DSEO en milligrammes par kilogramme de poids corporel plutôt qu'en milligrammes par kilogramme de nourriture, puisque les schémas diététiques varient d'une espèce animale à une autre et qu'ils diffèrent sensiblement en fonction des populations humaines. En second lieu, l'étude effectuée par la NAS reconnaissait de façon explicite que les résultats dose-effet obtenus directement chez l'homme ne nécessiteraient qu'un facteur de sécurité de 10 pour tenir compte des variations de sensibilité chez l'homme; par contre, les résultats d'expériences menées sur des animaux nécessiteraient un facteur de sécurité de 100 pour tenir compte des incertitudes additionnelles lors de l'extrapolation inter-espèces de l'animal à l'homme. Un autre facteur de sécurité de 10 devait y être inclus dans les cas où les données appropriées de toxicité chez l'homme et l'animal seraient absentes, portant la valeur du facteur de sécurité à 1000.

En bref, *Drinking Water and Health* a fait des recommandations explicites quant au type et à l'ordre de grandeur des facteurs de sécurité qui étaient semblables aux facteurs de sécurité de la FDA (NAS, 1977):

- Des résultats expérimentaux (épidémiologiques) valides d'études portant sur l'ingestion prolongée chez l'homme sans induction de carcinogénicité.

  (Facteur de sécurité = 10)

- Des résultats expérimentaux d'études de l'ingestion chez l'homme non disponibles ou insuffisants (par exemple, exposition aiguë seulement). Résultats valides d'études portant sur l'alimentation à long terme chez l'animal de laboratoire ou, en absence d'études menées chez l'homme, des études valables effectuées sur au moins une espèce animale. Pas d'indication de carcinogénicité.

  (Facteur de sécurité = 100)

- Aucun résultat d'études aiguës ou à long terme chez l'homme. Résultats insuffisants chez l'animal. Pas d'indication de carcinogénicité.

  (Facteur de sécurité = 1000)

À l'aide d'une méthodologie dérivée principalement des recommandations faites par la NAS en 1977, l'EPA a publié en 1980 *Guidelines and Methodology Used in the Preparation of Health Effects Assessment Chapters of the Consent Decree Water Criteria Documents* (EPA, 1980). Ce règlement présentait un ensemble préliminaire de critères pour le choix des paramètres dose-effet et des facteurs de sécurité appropriés, de même qu'une méthodologie pour la détermination de la DJA pour les toxiques à seuil présents dans l'eau. Les facteurs de sécurité spécifiés dans les règlements faits par l'EPA en 1980 concernant l'eau potable étaient essentiellement les mêmes que ceux qui avaient déjà été décrits dans *Drinking Water and Health*.

La méthodologie de la DJA a également servi à d'autres secteurs de réglementation de l'EPA, dont l'élaboration des normes pour la qualité de l'air ambiant et la gestion des déchets toxiques. Toutefois, l'absence de ligne directrice relative à l'utilisation de la DJA qui soit uniformisée dans les divers bureaux de réglementation de l'EPA a entraîné des incohérences graves dans la méthodologie de la détermination de la DJA et dans les vraies valeurs de DJA calculées par différents bureaux de l'EPA pour un même produit chimique:

> *...les bureaux chargés de programmes ont utilisé diverses méthodes pour évaluer le degré d'incertitude. Certains évaluent une DRéf (Dose de référence) chronique pour la vie (DJA), fondée sur les facteurs d'incertitude (facteurs de sécurité) reliés aux renseignements disponibles, et définissent alors les critères en fonction de cette DRéf. Certains calculent une marge de sécurité entre la plus haute dose qui n'entraîne pas d'effet délétère observable (DSEDO) de l'effet critique, et l'exposition estimée, et évaluent alors spécifiquement cette marge en fonction du produit chimique étudié et de son schéma d'exposition prévu. Certains bureaux évaluent un degré de protection approprié en considérant chaque cas individuellement, en utilisant leurs meilleures ressources techniques et scientifiques. Enfin, certains*

programmes ont mis au point leurs propres techniques quantitatives pour l'extrapolation ou l'interpolation de données manquantes; un petit nombre d'entre eux ont déjà reçu l'approbation générale au sein de l'Agence.

Preuss et Erlich (1987)

Pour cette raison, au cours des années quatre-vingt, l'EPA a tenté de préciser les définitions opérationnelles de différentes mesures dose-effet et de facteurs de sécurité, et de définir des méthodes de détermination de la DJA uniformisées, dans le but d'imposer un cohérence méthodologique dans tous les bureaux de l'EPA peu importe le mode d'exposition ou la juridiction en matière de réglementation. Par exemple, Dourson et Stara (1983) ont évalué systématiquement l'utilisation et la validité des facteurs de sécurité dont l'EPA se sert pour la détermination de la DJA. Plus tard, un grand nombre des facteurs de sécurité revus ont été dûment inclus dans les *National Primary Drinking Water Regulations* (EPA, 1985) pour la détermination des DJA relatives aux contaminants organiques et inorganiques présents dans les sources d'eau potable. Enfin, en 1988, l'EPA a publié des lignes directrices officiellement reconnues pour l'évaluation du risque relié à la présence de produits de toxicité systémique, sous le titre *Reference Dose (RfD): Description and Use in Health Risk Assessments* (Dose de référence [DRéf]: description et utilisation lors des évaluations du risque pour la santé) (Barnes et Dourson, 1988). Les EPA Guidelines de 1988 concernant les produits de toxicité systémique ont instauré une méthodologie uniformisée pour toutes les déterminations des DJA et des facteurs de sécurité effectuées par l'EPA pour les toxiques à seuil. Ces lignes directrices constituent un complément à la série des EPA Guidelines officielles déjà parue et portant sur l'évaluation du risque des produits cancérigènes, mutagènes et tératogènes.

B. Méthodologie relative au facteur de sécurité DSEO

1. Définition de la DJA

Dans le cas de produits toxiques présentant des comportements à seuil, il était d'usage courant de recourir à la méthodologie de la dose journalière admissible (DJA) pour déterminer des niveaux d'exposition «sûrs». Même si la méthodologie de la DJA a d'abord été mise au point pour l'établissement de normes de sécurité pour les additifs alimentaires et les résidus de pesticides, elle a par la suite été appliquée de façon générale à l'établissement d'un grand nombre de normes de tous genres, y compris celles relatives à l'exposition aux contaminants présents dans l'air et dans l'eau et aux déchets toxiques présents dans l'environnement (Dourson et Stara, 1983). La méthodologie de la DJA s'applique généralement aux expositions qui ont lieu pendant de longues périodes de temps (toxicité chronique) et vise à assurer que des expositions fréquentes ou continuelles ne causeront pas d'effets délétères pour la santé (Lu, 1983). De façon générale, la DJA se définit comme suit:

...la dose journalière d'un produit chimique (exprimée en milligrammes par kilogramme de poids corporel par jour) qui, au cours d'une vie entière, ne semble pas constituer un risque appréciable compte tenu de tous les faits connus à ce moment.

Lu (1983)

Par conséquent, cette définition pose un problème concernant surtout les effets sur la santé d'une exposition chronique à un produit au cours de la vie entière, et elle suggère que les normes de sécurité soient continuellement révisées à la lumière des plus récentes données toxicologiques. L'EPA a quant à elle formulé une définition opérationnelle plus explicite de la DJA:

La DJA se définit comme la quantité de produit toxique exprimée en milligrammes par kilogramme de poids corporel par jour... qui n'est pas censée causer d'effets délétères pour la santé par suite d'une exposition chronique pour la population humaine en général, y compris les sous-groupes de sensibilité. On considère comme effets délétères la détérioration fonctionnelle ou les lésions pathologiques qui peuvent affecter le bon fonctionnement de l'organisme entier, ou qui diminuent les facultés d'adaptation ou de défense d'un organisme.

EPA (1980)

Puisqu'il n'existe pas d'étude toxicologique menée sur un nombre restreint d'animaux de laboratoire et permettant d'assurer une sécurité absolue pour l'homme, la DJA doit être présentée comme un intervalle dont la limite inférieure est l'absence d'exposition, et dont la limite supérieure est l'exposition admissible à un produit chimique. Le JECFA a ainsi défini la DJA pour souligner le fait qu'elle représente une limite supérieure du risque admissible, et qu'il faut s'en tenir aux niveaux d'exposition les plus faibles que la technologie peut permettre (OMS, 1987). De la même façon, les EPA Guidelines de 1988 rejettent catégoriquement ce qui peut en dériver implicitement, soit que toute exposition à un niveau moindre que la DJA est «admissible» ou «sûre», ou que toute exposition à un niveau supérieur à la DJA est toujours «inadmissible» (Barnes et Dourson, 1988). La notion de DJA est définie comme suit:

En pratique, la DJA représente aux yeux d'un grand nombre de personnes (dont les gestionnaires de risque) un niveau admissible d'exposition et, par déduction, toute exposition qui excède la DJA est considérée comme «inadmissible». Cette démarcation rigide entre ce qui est «admissible» et ce qui ne l'est pas se heurte aux opinions des toxicologues, qui interprètent habituellement la DJA comme une estimation assez rudimentaire d'un niveau d'exposition chronique qui n'est pas susceptible de causer d'effets délétères chez l'homme. La DJA est généralement considérée par les personnes responsables de l'évaluation du risque

comme une estimation «non rigide», dont le degré d'incertitude peut couvrir un ordre de grandeur. Cela signifie que, toutes proportions gardées, même si l'exposition à des niveaux légèrement supérieurs à la DJA est associée à une probabilité plus élevée d'effets délétères, cette probabilité ne constitue pas une certitude. De la même façon, même si la DJA est considérée comme le niveau d'exposition auquel la probabilité d'effets délétères est faible, il est impossible de garantir l'absence de risque pour la population entière.

Barnes et Dourson (1988)

Pour cette raison, les EPA Guidelines de 1988 relatives aux produits de toxicité générale ont redéfini la DJA en termes de *dose de référence* (DRéf) pour éviter toute allusion à une «exposition admissible» ou à une «notion de sécurité absolue»:

...il ne faut pas conclure catégoriquement que toutes les doses inférieures à la DRéf sont «admissibles» (ou qu'elles ne représentent pas de risque), ou que toutes les doses supérieures à la DRéf sont «inadmissibles» (ou qu'elles causeront des effets délétères).

Barnes et Dourson (1988)

Les EPA Guidelines de 1988 définissent la DRéf comme suit:

De façon générale, la DRéf est une estimation (dont le degré d'incertitude peut couvrir un ordre de grandeur) de l'exposition quotidienne à laquelle peut être soumise la population humaine (y compris les sous-groupes de sensibilité) et qui est susceptible de ne pas présenter de risque appréciable d'effets délétères au cours de toute une vie.

Barnes et Dourson (1988)

Le calcul de la DRéf est analogue à celui de la DJA: la DRéf est le quotient de la dose sans effet délétère observé (DSEDO) par le facteur d'incertitude global (FI) et par un autre facteur de modification (FM).

$$DRéf = DSEDO / (FI \times FM)$$

Comme la méthodologie sous-jacente au calcul de la DJA et de la DRéf est essentiellement la même, les deux expressions seront considérées équivalentes dans cet article, sauf indications contraires.

2. Définition de la DSEO et d'autres estimateurs de seuil

Les études toxicologiques consistent normalement à exposer des groupes d'animaux de laboratoire à de multiples niveaux de dose d'un produit chimique. En pratique, chaque étude dose-effet doit comprendre au moins un niveau de dose qui n'entraîne aucun effet délétère, et qui représente la *dose sans effet observé* (DSEO). De plus, plusieurs niveaux de doses supérieurs qui entraînent des effets délétères doivent être inclus dans l'étude afin de définir la courbe dose-effet et ainsi démontrer la gravité des effets toxiques. Habituellement, la DSEO est établie d'après des animaux appartenant à des espèces appropriées les plus susceptibles, au moyen de l'indicateur de toxicité le plus sensible révélé au cours d'une étude de toxicité à long terme (chronique) (Lu, 1988).

Lorsque l'EPA a élaboré les lignes directrices pour l'établissement de normes d'exposition dans le cas des toxiques à seuil, elle a défini plusieurs types de niveaux d'effets (EPA, 1980; Dourson, 1986; Stara et coll., 1987). Dans l'ordre croissant de gravité, les niveaux d'effets sont les suivants:

- DSEO (dose sans effet observé)
  Dose qui n'entraîne pas d'effet observable. Le niveau d'exposition le plus élevé qui ne cause pas d'augmentation statistiquement ou biologiquement significative de fréquence ou de gravité des effets, par comparaison entre la population exposée et un groupe témoin approprié.

- DSEDO (dose sans effet délétère observé)
  Dose qui n'entraîne pas d'effet délétère observable. Le plus haut niveau d'exposition qui ne cause pas d'augmentation statistiquement ou biologiquement significative de la fréquence ou de la gravité des effets, par comparaison entre la population exposée et un groupe témoin approprié. *Il y a des effets à ce niveau, mais ils ne sont pas considérés comme délétères.*

- DMEO (dose minimale à effet observé)
  La plus faible dose qui entraîne des effets observables; la plus faible dose qui entraîne des effets délétères observables. Le plus faible niveau d'exposition qui cause une augmentation statistiquement ou biologiquement significative de la fréquence ou de la gravité des effets, par comparaison entre la population exposée et un groupe témoin approprié. *Si les effets qui peuvent apparaître à ce niveau ne sont pas considérés comme délétères, la DMEO équivaut à la DSEDO. Si les effets apparents à ce niveau sont considérés comme délétères, la DMEO équivaut à la DMEDO.*

- DMEDO (dose minimale à effets délétères observés)
  La plus faible dose qui entraîne des effets délétères observables. Le plus faible niveau d'exposition qui cause une augmentation statistiquement ou biologiquement significative de la fréquence ou de la gravité des effets, par comparaison entre la population exposée et un groupe témoin approprié. *Les effets apparaissant à ce niveau sont considérés comme délétères.*

- DEP (dose à effets patents)
  Dose qui entraîne des effets patents. Le niveau d'exposition qui entraîne des effets délétères patents, comme la détérioration fonctionnelle irréversible ou la mort, d'une façon statistiquement ou biologiquement plus marquée dans la population exposée que dans un groupe témoin approprié.

Puisque le chercheur doit déterminer si les modifications sont des manifestations de la toxicité ou si elles ne sont que des réactions d'adaptation à l'exposition, l'EPA a défini des termes spécifiques pour distinguer les niveaux d'effets dans le cas d'effets délétères graves, à savoir: la dose qui n'entraîne pas

d'effet délétère observable (DSEDO) et la plus faible dose qui entraîne des effets délétères observables (DMEDO). Selon les définitions de l'EPA (1980), la détermination de la DSEDO et de la DMEDO ne doit tenir compte que des effets toxiques véritables, sans les réponses physiologiques ou adaptatives, comme l'accroissement du volume hépatique par stimulation de la prolifération des microsomes ou l'amaigrissement par suite de l'absence de palatabilité des aliments, qui constitue un effet non pathogénique de l'exposition à un produit chimique:

> *On considère comme effets délétères la détérioration fonctionnelle ou les lésions pathologiques qui peuvent affecter le bon fonctionnement de l'organisme entier, ou qui diminuent les facultés d'adaptation ou de défense d'un organisme.*
>
> EPA (1980)

Selon le raisonnement de l'EPA, une DMEO peut en fait correspondre soit à une DSEDO, soit à une DMEDO: le jugement permet alors de savoir si la DMEO équivaut en fait à une DMEDO ou à une DSEDO. Toutefois, d'autres organismes de réglementation considèrent généralement les expressions DSEO et DMEO exclusivement comme des effets toxiques; ces expressions peuvent alors être utilisées comme des synonymes de DSEDO et de DMEDO. Par exemple, dans le cas d'ambiguïté quant aux effets toxiques relatifs aux additifs alimentaires, le JECFA suggère que *«dans le doute quant à la signification d'un effet particulier, il faut le considérer comme délétère»* (Lu,1988).

Plusieurs lignes directrices ont été formulées en fonction du système de classification dose-effet pour la dérivation des valeurs de DJA à partir des données dose-effet obtenues sur des animaux. En voici un résumé textuel:

1) *Une DEP qui n'est associée à aucune autre donnée ne peut pas servir à la détermination de la DJA.*

2) *Une DSEO qui n'est associée à aucune autre donnée ne peut pas servir à la définition des critères. En présence de plusieurs DSEO sans autres données relatives aux DMEO, DSEDO ou DMEDO, il faut utiliser la DSEO la plus élevée pour le calcul d'une DJA.*

3) *Une DSEDO, une DMEO ou une DMEDO peut être utilisée pour le calcul de la DJA. Une DSEDO convenablement définie d'après une étude chronique (d'au moins 90 jours) peut être utilisée directement, avec le facteur d'incertitude (sécurité) approprié. Il faut prendre garde de ne pas confondre les DEP et les DMEDO.*

4) *Si, dans le cas de doses raisonnablement rapprochées, seules une DSEO et une DMEO de qualité égale sont disponibles, le facteur d'incertitude (sécurité) approprié doit être appliqué à la DSEO.*

Source: Adaptation de EPA (1980).

Depuis 1980, l'EPA fait une distinction claire entre *dose qui n'entraîne pas d'effet* (DSE) en tant que mesure dose-effet servant à définir le seuil, et la *dose qui n'entraîne pas d'effet observable* (DSEO) en tant qu'estimateur statistique du seuil. Cette distinction indique que la DSEO ne constitue pas une dose absolue qui n'entraîne pas d'effet, mais plutôt une approximation expérimentale de la DSE soumise à l'influence de la taille de l'échantillon et à la sensibilité de l'espèce étudiée (EPA, 1980; Nelson, 1984; Crump, 1984). La DSE correspond alors à un niveau posologique situé quelque part entre la DSEO et la DMEO, malgré l'imprécision sur la valeur réelle de la DSE. Dans le cas de doses très espacées, la DSEO peut se trouver très en-dessous de la valeur réelle de la DSE, et la DMEO, très au-dessus.

La DSEO ne constitue qu'une estimation de la DSE réelle, car elle dérive d'un échantillon restreint d'animaux qui ont été observés selon un schéma arbitraire de conditions expérimentales. En utilisant d'autres conditions expérimentales avec une espèce plus susceptible ou un indicateur de toxicité plus sensible, on peut s'attendre à trouver une valeur plus faible pour la DSEO.

> *Une DSEO n'est pas une propriété inhérente à un système animal, mais dépend plutôt du schéma expérimental et des résultats. Elle représente alors, en termes statistiques, une estimation ou une valeur statistique de la «dose réelle qui n'entraîne pas d'effet». Cette dernière expression représente la dose la plus élevée qui reste absolument sûre: c'est une propriété inhérente au système, ou un paramètre, en termes statistiques. Dans le cas d'un effet pour lequel il n'y a pas de seuil, la «dose réelle qui n'entraîne pas d'effet» est nulle.*
>
> Crump (1984)

De la même façon, la valeur estimée de la DSEO est déterminée en partie par la sensibilité statistique de l'expérience dose-effet, qui est fonction de la taille de l'échantillon des animaux exposés au produit étudié. Par conséquent, la détermination d'une DSEO dans un système animal ne prouve pas vraiment l'existence d'un seuil:

> *L'existence d'une DSEO ne prouve ni ne réfute l'existence ou le niveau d'un seuil réel; elle ne définit que le niveau d'exposition le plus élevé qui, dans des conditions expérimentales données, n'est pas associé à une augmentation significative d'effet.*
>
> IRLG (1986)

Il existe alors une certaine limite à l'utilité statistique de ces définitions, car la pente et l'allure globale de la courbe dose-effet, ainsi que l'effet de la taille de l'échantillon sur les intervalles de confiance statistiques n'entrent habituellement pas en ligne de compte dans la détermination de la DSEO (Munro et Krewski, 1981; Crump, 1984; Krewski et coll., 1984). D'autres méthodes d'estimation de la DJA, basées sur l'analyse statistique de toute la courbe dose-effet ont été proposées pour le traitement des données provenant d'expériences effectuées sur des animaux; leur application pratique n'est toutefois pas encore très répandue (Gaylor, 1983; Crump, 1984; Dourson, 1986).

C. Description des facteurs de sécurité

1. Définition d'un facteur de sécurité

La U.S. National Academy of Science a donné une définition fondamentale d'un facteur de sécurité:

> *«Facteur de sécurité» ou «facteur d'incertitude» désigne un nombre qui reflète le degré d'incertitude qui doit être considéré au moment de l'extrapolation de données expérimentales de l'animal à l'homme.*
>
> NAS (1977)

Dernièrement, l'EPA a introduit le terme «facteur d'incertitude» (FI) pour remplacer le facteur de sécurité (FS), afin d'éviter les glissements vers la notion de sécurité absolue associée aux valeurs de DJA calculées par l'application des facteurs de sécurité aux DSEO tirées d'expériences effectuées sur des animaux (EPA, 1980, 1988; Dourson et Stara, 1983):

> *Le terme «facteur de sécurité» suggère, peut-être malgré lui, la notion de sécurité absolue, c'est-à-dire, l'absence de risque. En dépit de fondements théoriques à l'appui de l'existence d'un seuil et d'une «sécurité absolue» reliés à certains produits chimiques, les preuves expérimentales évidentes sont absentes dans la plupart des cas.*
>
> Barnes et Dourson (1988)

Pour plus d'uniformité, le terme facteur de sécurité (FS) sera considéré comme synonyme de facteur d'incertitude (FI).

2. Choix des facteurs de sécurité

D'après les EPA Guidelines de 1988 (Barnes et Dourson, 1988), quatre facteurs de sécurité (incertitude) normalisés sont définis en bloc (Tableau 2.1), chacun possédant une valeur par défaut fixée à 10:

- **variabilité humaine** (FI = 10H)
- **extrapolation inter-espèces** (FI = 10A)
- **correction pour études subchroniques** (FI = 10S)
- **correction pour l'utilisation de la DMEO** (FI = 10L)

De cette façon, le facteur de sécurité traditionnel 100 a été formellement décomposé en ses deux composantes principales. Le premier facteur normalisé de sécurité de 10 est introduit pour tenir compte de la variation individuelle de la sensibilité de l'homme aux produits chimiques (10H). Le second facteur normalisé de sécurité de 10 tient compte de l'incertitude quant à l'extrapolation inter-espèces entre l'animal et l'homme (10A). Pour compenser l'utilisation d'une DSEO obtenue d'une étude subchronique (90 jours) effectuée sur l'animal, en l'absence de données comparables tirées d'une étude chronique (180-360 jours), on doit inclure un autre facteur normalisé d'incertitude de 10 pour les études subchroniques (10S). Enfin, s'il n'est pas possible de déterminer directement la DSEO à partir des données dose-effet parce qu'elle se situe en-deçà des concentrations normalement utilisées au cours des expériences animales, il faut alors substituer la DMEO à la DSEO en incluant un autre facteur de sécurité de 10 pour obtenir une valeur de DMEO de l'ordre de celle de la DSEO (10L).

Tableau 2.1
Lignes directrices et preuves expérimentales en faveur de l'utilisation de facteurs d'incertitude (sécurité)*

(adaptation de Dourson et Stara, 1983)

Lignes directrices	Preuves expérimentales
Utiliser un facteur 10 pour l'extrapolation de résultats expérimentaux valides tirés d'études portant sur l'ingestion prolongée chez l'homme. Ce facteur 10 protège les personnes sensibles dans la population, d'après une estimation des résultats recueillis auprès de personnes moyennes en bonne santé.	Analyse log-probit; analyse log-probit; sensibilité humaine composite.
Utiliser un facteur 100 pour l'extrapolation de résultats valides tirés d'études d'alimentation à long terme menées sur des animaux de laboratoire et de résultats d'études portant sur l'ingestion chez l'homme non disponibles ou in complets (par exemple exposition aiguë seulement). Il s'agit ici d'un autre facteur 10 d'incertitude pour tenir compte de l'extrapolation à l'homme moyen des données obtenues à partir de l'animal moyen.	Équivalence de dose relative à la surface corporelle; comparaison de la toxicité d'une espèce à une autre.
Utiliser un facteur 1000 pour l'extrapolation de résultats tirés d'expériences de plus courte durée (subchroniques) effectuées sur des animaux de laboratoire en absence de données utilisables d'expériences à long terme ou aiguës menées chez l'homme. Il s'agit ici d'un autre facteur 10 d'incertitude pour tenir compte de l'extrapolation de données tirées d'expérience variant de subchroniques à chroniques.	Comparaison de la DSEO subchronique/chronique; comparaison de la DSEO ou de la DMEO subchronique/chronique.
Utiliser un facteur d'incertitude additionnel valant entre 1 et 10 selon la sensibilité de l'effet délétère pour la dérivation d'une DJA à partir d'une DMEO. Ce facteur d'incertitude rabaisse la valeur de la DMEO à une valeur proche de celle de la DSEO.	Comparaison DMEO/DSEO.

*Ces facteurs doivent être appliqués à la DSEDO ou à la DSEO valide la plus élevée qui n'a pas de DMEO qui lui est inférieure ou égale, pour le calcul d'une DJA dans le cas où il n'existe aucune indication sur la nature cancérigène d'un produit chimique.

Dans une situation simple, le facteur de sécurité global peut être calculé par la multiplication des facteurs de sécurité normalisés, choisis en fonction des critères décrits au Tableau 2.1. Ensemble, les facteurs de sécurité normalisés produisent un facteur de sécurité globlal dont l'ordre de grandeur est de 10, 100 ou 1000:

$$FS_{TOTAL} = \prod_{i=1}^{n} FS_i = FS_1 \times FS_2 \times FS_3 \times FS_4$$

Les EPA Guidelines de 1988 prévoient également l'inclusion d'un *facteur de modification (FM)* servant de facteur de correction subjectif et tenant compte de tout renseignement additionnel qui peut affecter la valeur de la DJA, établi «au mieux du jugement» des évaluateurs du risque. La raison d'être de l'inclusion de ce facteur de modification est la suivante:

> *Les facteurs de sécurité ne prétendent pas avoir une précision mathématique. Lorsque les renseignements se font rares, les facteurs de sécurité restent la meilleure solution pour évaluer le risque pour les faibles niveaux d'exposition. L'utilisation d'un modèle mathématique complexe dans de telles situations peut donner une illusion trompeuse de précision. Par contre, lorsque les renseignements sont plus abondants, par exemple, dans le cas où il existe un mécanisme connu de bio-accumulation ou encore des méthodes de dosage d'enzymes permettant de suivre le métabolisme de l'enzyme, je crois qu'il est alors de première importance de tenir compte de ces renseignements dans le processus d'évaluation. Dans de telles situations, il serait incorrect de ne pas tenir compte de ces données et d'utiliser une méthode simple qui consiste à choisir bêtement une certaine puissance de dix en guise de facteur de sécurité convenable.*
>
> EPA Science Advisory Board (EPA, 1985)

Le FM peut prendre une valeur supérieure à zéro jusqu'à un maximum de 10; sa valeur par défaut est fixée à l'unité. Il est ainsi possible d'obtenir un facteur de sécurité global corrigé qui n'est pas toujours une puissance de 10. Par exemple, le facteur de correction subjectif (FM) d'une étude dose-effet effectuée sur un échantillon de grande taille (250 animaux) pourrait être de 0,8. Le produit de ce FM par un FI courant de 1000 donnerait alors un facteur de sécurité global corrigé de 800 (Barnes et Dourson, 1988).

D'après les EPA Guidelines de 1988, les principales corrections apportées par le FM tiennent compte des renseignements manquants dans la base de données et du nombre d'espèces soumises aux expériences. Elles ne servent toutefois pas à définir des critères décisionnels ou à spécifier la dimension des facteurs de correction. Certains critères additionnels concernant l'application d'un facteur de modification ont déjà été décrits dans les *National Primary Drinking Water Regulations* (EPA, 1985):

- exposition d'animaux de laboratoire pendant une partie seulement de leur vie
- étude effectuée au cours d'une partie de la vie des sujets (vraisemblablement pour un suivi pathologique)
- étude portant sur une autre voie (mode) d'exposition (par exemple nourriture plutôt que l'eau)
- le produit chimique est un élément nutritif essentiel à faible dose, mais toxique à dose élevée
- l'effet était beaucoup plus sensible que les effets traditionnels utilisés pour les calculs de la DJA.

Les trois premiers critères énoncés plus haut donneraient lieu à un FM supérieur à 1, augmentant ainsi la valeur du facteur de sécurité corrigé, et diminuant par le fait même la valeur de la DJA. Les deux derniers critères produiraient un FM inférieur à 1, donnant ainsi un facteur de sécurité corrigé plus petit et une valeur de DJA supérieure à celle obtenue autrement.

De tels critères permettent l'obtention de valeurs de *facteurs de sécurité corrigés* inférieures à 10. Par exemple, dans le cas de produits chimiques dont l'effet est bien documenté chez des personnes sensibles, comme les suppléments de fluorure dans l'eau potable, un FM de 0,1 pourrait s'appliquer à un facteur de sécurité déjà établi à 10, produisant ainsi un facteur de sécurité corrigé égal à un seulement (Barnes et Dourson, 1988). Cet exemple peut cependant se révéler trompeur, puisque l'effet choisi dans le cas du fluorure est l'aspect tacheté des dents, qui ne constitue qu'un effet esthétique en réalité sans danger. La dose minimale de fluorure pouvant causer des effets nocifs pour la santé se situe dans une échelle beaucoup plus grande.

Lorsque les données ne cadraient pas précisément avec les exigences d'une catégorie de facteurs de sécurité, l'EPA avait l'habitude de permettre la modification de la dimension d'un facteur de sécurité global à l'aide de *facteurs intermédiaires de sécurité*, établis d'après des critères non normalisés comme la gravité des points critiques de toxicité et les caractéristiques dose-effet du produit étudié (EPA, 1980). Un tel facteur intermédiaire de sécurité pouvait avoir une valeur logarithmique (par exemple 32 se situe à mi-chemin entre 10 et 100 sur une échelle logarithmique). De plus, le facteur de sécurité DMEO pouvait adopter n'importe quelle valeur entre un et 10, qui relève de décisions subjectives, comme la valeur absolue plus ou moins grande de la pente de la courbe dose-effet. (EPA, 1985). Les produits dont la courbe dose-effet possède une pente de valeur absolue élevée sont plus susceptibles d'avoir une DMEO semblable à la DSEO que les produits dont la courbe possède une pente moins prononcée. Lorsque la valeur de la DJA est tirée de la DMEO, le facteur global de sécurité se trouverait alors entre 100 et 1000, ce qui exigerait un facteur de correction d'autant plus petit que la pente des courbes est plus prononcée.

Cependant, les plus récentes EPA Guidelines (1988) ont inclus dans la valeur du facteur de modification ces corrections subjectives, au lieu d'ajuster la valeur de chaque facteur de sécurité comme cela a été décrit plus haut. Néanmoins, les EPA Guidelines de 1988 ont évité avec grand soin de fournir des

critères décisionnels détaillés pour la correction des FM selon des considérations scientifiques subjectives. Une telle énumération de critères pourrait par exemple comprendre: l'analyse de la variabilité statistique de la DSEO et des FI sous-jacents, la caractérisation de la courbe dose-effet (y compris la pente), et la gravité des effets biologiques. Présentement, le personnel scientifique de l'EPA investit beaucoup d'efforts dans la mise au point de l'inclusion officielle de ces critères pour le calcul du FM et du facteur corrigé de sécurité.

D. Validation des facteurs de sécurité

1. Facteur global de sécurité

E.J. Bigwood, travaillant pour le compte de l'OMS (Organisation mondiale de la santé), a écrit un article de revue (1973) dans lequel il expliquait en détail le raisonnement qui sous-tendait l'utilisation du facteur de sécurité de 100, au moyen de l'analyse de diverses sources de variation inter-espèces, et par la segmentation du facteur de sécurité de 100 réparti entre ces sources. Les cinq sources de variation ont été identifiées comme suit:

- taille corporelle (animal de laboratoire et homme)
- schémas diététiques
- équilibre hydrique (consommation)
- fonctions hormonales et leur effet sur l'apport alimentaire
- différences inhérentes de sensibilité aux produits toxiques.

Des valeurs approchées, estimées selon les données toxicologiques existantes, ont été arbitrairement attribuées à chacune des quatre premières catégories: un facteur de sécurité d'environ 60 s'avérait nécessaire pour tenir compte des quatre premières sources de variation. Quant à la cinquième source, qui était difficilement quantifiable, l'équipe de E.J. Bigwood a suggéré qu'il était raisonnable de tenir compte des cinq sources de variation en faisant passer la valeur du facteur de sécurité de 60 à 100. En conséquence, on a conclu que la valeur de 100 pour un facteur de sécurité était hautement valide et que les déviations au-delà ou en-deçà de cette valeur (par exemple de 50 à 500) n'étaient pas nécessaires dans la plupart des cas. E.J. Bigwood concluait que le facteur de sécurité de 100 ne requiert des corrections que dans des situations extraordinaires.

2. Facteurs spécifiques de sécurité

Dans le but de fournir des preuves théoriques et empiriques en faveur du bien-fondé de chacun des facteurs de sécurité de 10 recommandés par les EPA Guidelines de 1988, Dourson et Stara (1983) ont entrepris une étude systématique de validation à l'aide de bases de données déjà existantes concernant la toxicité chez l'animal.

a) Variation intra-espèce

Le facteur de sécurité de 10 tenant compte des variations individuelles de sensibilité dans une espèce donnée a été validée au moyen d'une analyse des données dose-effet tirées de 490 pentes distinctes de courbes probit/log dose provenant d'expériences de létalité aiguë (DL_{50}). Un sous-groupe sensible de la population a été défini par la réduction de la réponse médiane (DL_{50}) de 3 probit, ce qui correspondait à un effet cumulatif de 0,13% ($DL_{0,13}$). La pente de la courbe (probit/log dose) d'extrapolation tracée entre les points DL_{50} et $DL_{0,13}$ a alors servi de mesure de la variation individuelle de sensibilité entre les individus dans la moyenne (médiane) et les individus les plus sensibles d'une population.

Les courbes de pente prononcée (>3) indiquent qu'un facteur de sécurité de 10 réduirait la dose appliquée à un niveau inférieur à la $DL_{0,13}$, protégeant ainsi convenablement les sous-groupes plus sensibles de la population. Les courbes de pente moins prononcée (<3) indiquent qu'un facteur de sécurité de 10 ne suffirait pas à réduire la dose appliquée sous la $DL_{0,13}$, ce qui par conséquent ne protégerait pas les individus les plus sensibles de la population. En fait, 92% des pentes des courbes probit/log dose avaient une valeur supérieure à trois, ce qui démontre que le facteur de sécurité de 10 est suffisant pour tenir compte de la variation de sensibilité à la plupart des produits chimiques dans une espèce donnée.

Les résultats de cette étude sont certes rassurants, mais les résultats d'une étude de validation intra-espèce menée sur des animaux ne peuvent être appliqués directement à l'homme. La faible variation d'effets observée chez l'animal reflète l'homogénéité génétique et la stabilité des conditions de vie controlées de l'animal de laboratoire. Comme la population humaine est génétiquement hétérogène, et que les milieux de vie des humains varient plus que ceux des animaux de laboratoire, il est probable d'observer une plus grande variation de la sensibilité aux produits chimiques chez l'homme que chez l'animal.

Calabrese (1985) a relevé d'importantes variations interindividuelles chez l'homme au sujet de la pharmacocinétique des produits chimiques présents dans l'environnement et du processus de réparation de l'ADN à la suite d'une atteinte possible par un ou plusieurs produits. Même si la majorité des personnes étudiées réagissaient dans l'intervalle d'un facteur de 10, la variation observée chez les personnes les plus sensibles dépassait souvent un facteur de 10, qui atteignait quelquefois 1000. Selon le produit chimique auquel est exposée une population, le facteur de sécurité de 10 tenant compte de la variabilité humaine semble fournir une protection adéquate pour environ 80-95% de la population; les autres 5-20% réagissent selon un facteur supérieur à 10. De plus, puisque la variation due à chaque facteur pharmacocinétique a été évaluée individuellement, les effets combinés de plusieurs polymorphismes pharmacocinétiques chez des personnes hypersensibles pourraient se révéler beaucoup plus importants que ce qu'avaient prévu les études des facteurs individuels (Calabrese, 1978b).

Hattis et coll. (1987) ont rapporté des résultats similaires dans une analyse biostatistique de la variabilité humaine, au moyen de mesures de la variabilité interindividuelle pour différents paramètres pharmacocinétiques, dont la demi-vie du produit dans l'organisme, les concentrations sanguines maximales, et l'aire sous la courbe (ASC). À partir de données relatives à 101 produits chimiques, une analyse secondaire de

la distribution statistique (semi-logarithmique) de ces mesures a révélé que 10 constituait une valeur par défaut appropriée pour le facteur de sécurité; toutefois, dans certains cas particuliers, cette valeur pourrait ne plus être adéquate. Aussi, ces résultats ne tiennent compte que d'un aspect de la variation interindividuelle de sensibilité chez l'homme, soit la variabilité pharmacocinétique. D'autres facteurs, comme les variations des niveaux d'exposition et de la différence des réactions toxicodynamiques à des produits toxiques n'ont pas été pris en considération. Un fait reste préoccupant: une grande part du facteur de sécurité pour la variation humaine ne rend compte que des différences dans les schémas pharmacocinétiques.

b) Extrapolation inter-espèces

Puisque l'animal de laboratoire est de plus petite taille que l'homme, un facteur de conversion pour la taille corporelle devient nécessaire au moment de la transformation de la DJA établie d'après des expériences effectuées sur l'animal à une valeur adaptée pour l'homme (Dourson et Stara, 1983). Plusieurs facteurs d'échelle peuvent alors servir de base pour cette transformation; cependant, la surface corporelle constitue souvent un meilleur paramètre que le poids corporel pour la conversion inter-espèces, parce qu'elle tient mieux compte de la variation de taux de métabolisme d'une espèce à une autre (Calabrese, 1983, 1987; Vocci et Farber, 1988). Il est alors admis que différentes espèces ont la même sensibilité à une dose donnée par unité de surface corporelle, parce que cette méthode tient compte des taux relatifs de détoxication ou d'élimination des produits chimiques. Le calcul indirect de la surface corporelle se fait habituellement en portant le poids corporel à la puissance 2/3. Le *facteur de correction inter-espèces* générique pour mettre à l'échelle correspond alors à la racine cubique du rapport entre le poids du corps humain (70 kg) et le poids W de l'animal:

$$\text{Facteur de correction inter-espèces} = (70/W)^{1/3}$$

Ce facteur de conversion est censé produire une dose équitoxique chez l'animal de laboratoire et chez l'homme. La dose humaine est le quotient de la dose administrée à l'animal (mg/kg poids corporel) par le facteur de correction inter-espèces. L'application de la formule d'échelle aux données tirées d'études de sensibilité inter-espèces conduit à un facteur de correction inter-espèces situé entre six et 12, ce qui correspond aux résultats de l'étude de la NAS sur les différences inter-espèces de sensibilité (NAS, 1977).

Plusieurs études empiriques de changement d'échelle sont mentionnées en faveur du fait que la valeur du facteur de correction inter-espèces se situe généralement autour de 10: cette valeur est la même que celle du facteur de sécurité utilisé pour l'extrapolation inter-espèces (Dourson et Stara, 1983). Toutefois, un grand nombre de ces études sont de pertinence et de validité douteuses: par exemple, les études comparatives portant sur la sensibilité de l'animal et de l'homme pour une série de médicaments alkylants anti-tumoraux. Il est également inquiétant de constater qu'une proportion importante du facteur de sécurité de 10 pour l'extrapolation inter-espèces ne tient compte que des différences de taille corporelle. Par exemple, Dourson et Stara évaluent à environ six le facteur de correction rat-homme. Il reste alors une étroite marge pour tenir compte des autres facteurs inter-espèces, comme les différences dues aux facteurs pharmacocinétiques et à la sensibilité propre à chaque espèce, qui ne dépendent pas de la taille corporelle.

De plus, il existe de profonds désaccords entre les différents organismes de réglementation quant à la notion de surface corporelle. La correction pour la surface corporelle a été examinée, puis rejetée, par la JECFA et le JMPR, pour les raisons suivantes: le métabolisme des xénobiotiques ne présente pas nécessairement une bonne corrélation avec les taux métaboliques généraux de l'organisme, et la conversion des surfaces corporelles n'est pas étayée par des résultats expérimentaux (Lu 1985, 1988). Dans le cas de 28 pesticides organophosphorés dont les données dose-effet sont bien caractérisées, il a été démontré que la conversion des surfaces corporelles ne rend pas bien compte des différences de sensibilité entre le chien et le rat. Lu conclut que la conversion en fonction des surfaces corporelles ne doit pas être considérée comme meilleure que celle des poids corporels pour l'extrapolation inter-espèces.

Récemment encore, l'EPA n'utilisait le facteur de conversion pour les surfaces corporelles que pour l'estimation du risque concernant l'exposition aux produits cancérigènes pendant toute une vie. Maintenant, l'EPA songe à utiliser cette technique à la fois pour les produits cancérigènes et pour les produits de toxicité systémique, même si aucune ligne directrice officielle n'a encore été publiée (Stara et coll., 1987). Les techniques de calcul de la DJA utilisées dans d'autres pays (par exemple, le Conseil néerlandais de la santé, Dutch Health Council) a permis une diminution du facteur de sécurité de 100 fois à 30 fois, dans le cas des produits de toxicité systémique, lorsque la conversion en fonction des surfaces corporelles est comprise dans les calculs de la DJA (Vocci et Farber, 1988). Toutefois, comme le désaccord persiste au sujet de la conversion en fonction des surfaces corporelles, une telle modification de la dimension du facteur de sécurité est discutable.

c) Correction subchronique/chronique

Afin de déterminer l'effet de l'utilisation de données tirées d'études dose-effet subchroniques pour le calcul des valeurs de DSEO, Dourson et Stara (1983) ont comparé les DSEO obtenues d'études (2 ans) dose-effet chroniques à ceux obtenus d'études subchroniques dont la durée s'étendait de 30 à 310 jours (moyenne: 92 jours). De la même manière, les valeurs de DMEO ont été comparées. Les valeurs de DSEO ou de DMEO pour divers produits chimiques ont été tirées d'une série d'études toxicologiques menées par Weil et McCollister (1963). Le rapport entre la DSEO (ou la DMEO) subchronique et la DSEO (ou la DMEO) chronique témoigne du degré de surestimation de la valeur de la DJA. Des comparaisons de 30 études avec données de DSEO et de 22 études avec données de DMEO ont permis d'établir que la différence entre les valeurs de DSEO ou de DMEO tirées d'études chroniques et celles provenant d'études subchroniques est inférieure à deux dans plus de la moitié des cas. Dans 96% des études, les rapports de DSEO subchronique/chronique étaient inférieurs à 10. Ces résultats sont à l'appui de l'utilisation de 10 comme

facteur de sécurité pour tenir compte, avec un degré de confiance assez élevé, des incertitudes générées par l'estimation des DSEO d'après des études subchroniques. Toutefois, la possibilité qu'un changement pathologique infime ou latent passe inaperçu lors d'études subchroniques ne peut pas être entièrement exclue.

d) Correction DMEO/DSEO

Dourson et Stara (1983) ont également étudié les effets de l'utilisation de la DMEO au lieu de la DSEO pour le calcul de la DJA, à partir des études toxicologiques décrites plus haut (Weil et McCollister,1963). Pour chaque étude subchronique ou chronique, le rapport DMEO/DSEO peut être considéré comme la diminution de la valeur de la DMEO qui est nécessaire pour obtenir une valeur de la DSEO correspondante. Le rapport mesure alors le degré de surestimation de la valeur de la DJA, lorsque le calcul est fait avec les DMEO au lieu des DSEO. Les résultats ont montré que pratiquement tous les produits chimiques étudiés présentaient un rapport DMEO/DSEO inférieur à 10; dans 96% des cas, le rapport était inférieur à 5. Ces résultats appuient ainsi la thèse voulant qu'un facteur de sécurité de 10 suffise pour justifier l'utilisation des valeurs de DMEO, quand les valeurs de DSEO ne sont pas disponibles pour un produit chimique donné. Cette affirmation ne vaut toutefois que dans le cas d'expériences à doses multiples qui sont bien conçues et pour lesquelles on peut être assuré que la DMEO se situe dans le même intervalle de doses que la DSEO.

3. Autres conceptions des méthodes FS

A. Critiques de la méthode DSEO-FS

Des techniques statistiques très complexes ont été mises au point pour l'évaluation quantitative du risque présenté par les produits cancérigènes: en comparaison, la caractérisation de la relation dose-effet dans les méthodes DSEO-FS traditionnelles est très limitée. Par conséquent, plusieurs faiblesses méthodologiques ont été reprochées à la méthode DSEO-FS: un grand nombre de statisticiens considèrent que ces faiblesses compromettent la précision et l'utilité de l'évaluation de la DJA (Munro et Krewski, 1981; Crump, 1984). Parmi ces faiblesses, on compte:

- utilisation inadéquate du point DSEO

D'une façon générale, la valeur de la DSEO n'est considérée comme utile que lorsqu'elle représente un des nombreux points expérimentaux d'une courbe dose-effet bien caractérisée (EPA, 1980). Une DSEO expérimentale qui est isolée ne peut être d'aucune utilité puisqu'il n'y a pas d'indication quant à sa proximité du DMEO. Une DSEO qui est isolée peut alors se situer à plusieurs ordres de grandeur sous la région seuil. Il est également possible de sous-estimer considérablement la DSEO si les points expérimentaux sont rares et très distancés les uns des autres. Il s'avère donc nécessaire de disposer d'un nombre adéquat de points (doses) assez rapprochés (en général, au moins 3 à 5 points) pour obtenir une détermination plus précise de la DJA à partir de la DSEO.

- effets de la taille de l'échantillon

La sensibilité statistique d'une expérience dose-effet dépend principalement du nombre d'animaux utilisés dans l'étude. Une étude portant sur un échantillon de grande taille offre plus de possibilités de détecter un résultat statistiquement significatif que l'étude portant sur un échantillon de taille plus restreinte. Différentes interprétations statistiques sont évidemment possibles pour des taux de 0/10 et 0/1000, même si en principe ces deux taux peuvent servir à déterminer une DSEO (Munro et Krewski, 1981). Dans le cas des produits cancérigènes, l'évaluation quantitative du risque présente souvent des limites supérieures d'intervalles de confiance de 95% et servent d'estimation prudente de la limite supérieure du risque: à l'opposé, la DSEO est une valeur statistique qui constitue la meilleure estimation du risque sans qu'il ne soit jamais question de la limite supérieure du risque (Crump, 1984). Par conséquent, les DSEO et les valeurs de DJA qui en découlent seront plus petites et plus restrictives dans les expériences à grande échelle; et les valeurs de DJA obtenues à partir d'expériences effectuées à plus petite échelle seront plus grandes et moins restrictives. Il est donc encourageant d'entreprendre des études portant sur un échantillon de grande taille, puisque ces études sont plus susceptibles de produire de plus faibles estimations de DSEO et des restrictions plus sévères quant à la DJA d'exposition. Des tailles minimales d'échantillons doivent donc être spécifiées au moment de la planification d'expériences dose-effet, pour que les DSEO et les DJA ne soient pas trop surestimés.

- caractérisation de la courbe dose-effet

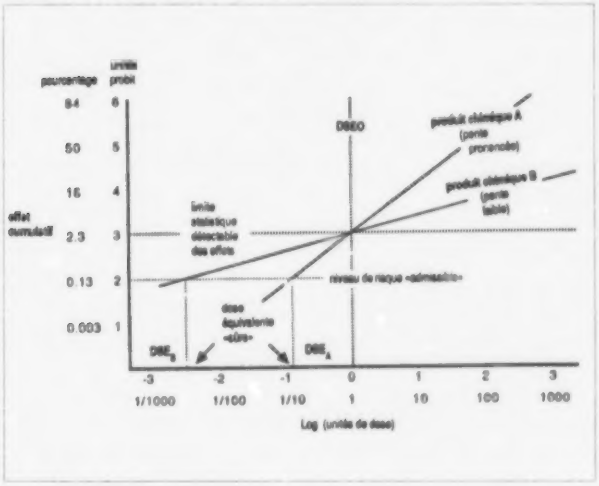
Lorsqu'on aborde la DSEO-FS de la façon traditionnelle, on n'utilise qu'un seul point expérimental de la courbe dose-effet pour calculer la DJA: la DSEO, ou sinon, la DMEO (Crump, 1984). Les autres points de la courbe dose-effet ne servent qu'à assurer une quantité et une répartition adéquate des points de façon à fournir une estimation raisonnable de la DSEO. La méthode d'estimation de la DSEO-FS ne tient donc habituellement pas compte de la pente et de l'allure de la courbe dose-effet (Munro et Krewski, 1981; Crump, 1984). Il y a par conséquent une perte de renseignements au sujet de la partie inférieure globale de la courbe, et il devient impossible de donner une estimation plus précise de la «véritable dose sans effet» (DSE), ou encore de tracer une courbe d'extrapolation pour les faibles doses afin d'estimer la dose-repère qui correspond à un niveau de risque prédéfini.

Dans le cas d'une courbe à pente prononcée, l'extrapolation de la courbe dose-effet vers le niveau d'effet inexistant est brève: la pente permet alors de prévoir une DSE relativement élevée. À l'opposé, une courbe à pente plus faible doit être extrapolée plus «longuement» vers l'effet nul, de sorte que la valeur de la DSE apparente est plus petite. Sur les courbes à pente faible, la distance entre les points DSEO et DMEO est plus grande que sur les courbes à pente plus

prononcée: par conséquent, la dose seuil pourrait être considérablement sous-estimée dans le cas des courbes à pente faible lorsque seule la DSEO est présente.

Par exemple, deux produits chimiques (*A* et *B*) possèdent des DSEO apparemment identiques, mais présentent des courbes dose-effet de pentes différentes: leurs doses-repères respectives à l'intérieur de la région inférieure de la courbe seront très différentes (Figure 3.1). Supposons qu'un niveau d'effet de 2,3% (équivalent à trois unités probit sur un tracé probit/log dose se situe juste sous la limite de détection statistiquement significative pour les deux produits, la DSEO est dérivée pour A et B comme se trouvant à un unité de dose de l'abscisse du point correspondant à un niveau d'effet de 2,3% (3 probits) sur la courbe. Si toutefois le niveau «admissible» de risque est défini à 0,13% (2 probits), ce qui représente une DSE pour 99,87% de la population, alors la dose équivalente «sûre» (dose-repère) serait environ le dixième d'une unité de dose pour A (pente prononcée), mais le 1/300e d'une unité de dose pour B (pente faible). Leurs DSE respectives (qui ne paraissent pas sur un tel tracé) seraient aussi très divergentes.

Figure 3.1
Effet de la pente sur le risque à faible dose



- valeurs arbitraires pour le facteur de sécurité

Les valeurs présentement attribuées aux différents facteurs de sécurité proviennent en fait d'une combinaison accident historique-évaluation rationnelle des sources d'incertitude pour le calcul des DJA. «*Le fait que l'homme possède dix doigts a sans l'ombre d'un doute joué un rôle dans le choix du facteur spécifique.*» (Crump, 1984). Les EPA Guidelines de 1988 qui définissent la dimension et l'utilisation des facteurs de sécurité normalisés ne fournissent pas d'explication sur la raison d'être ou la logique de la correction de ces facteurs en fonction de la taille de l'échantillon, du nombre et de la répartition des points expérimentaux, ou de la pente de la courbe dose-effet; ces critères figurent toutefois timidement dans la liste de critères nécessaires à l'établissement de la valeur du facteur de modification (FM). De plus, la méthode DSEO-FS ne fournit aucune technique pour prévoir les niveaux de risque aux doses inférieures à la DSEO, ou pour calculer les doses-repères correspondant aux niveaux de risque prédéfinis. Pour éliminer quelques-uns de ces sujets de critique, qui touchent essentiellement la caractérisation statistique de la courbe dose-effet, des chercheurs ont récemment proposé un certain nombre de nouveaux modèles pour les facteurs de sécurité, modèles fondés principalement sur l'analogie avec des modèles mathématiques d'ajustement de courbe déjà mis au point pour l'évaluation quantitative du risque présenté par les produits cancérigènes.

B. Modèles mathématiques pour le facteur de sécurité

1. Modèles de distribution normale (probit)

Les premières tentatives pour tenir compte de la pente de la courbe dose-effet dans la détermination de la DJA étaient basées sur les propriétés statistiques de la distribution log-normale de la relation dose-effet rencontrée dans la plupart des expériences effectuées sur des animaux. En 1956, on proposait déjà d'obtenir un facteur de sécurité en mesurant la pente de la courbe log dose-effet et son écart-type. La DJA pouvait être dérivée par la détermination de la dose seuil (c'est-à-dire la DSEO) et par l'extrapolation vers le bas à six écarts-types (Vettorazzi, 1976). En effet, on calculait la DJA en retranchant six unités probit de la DSEO, la soustraction constituant le facteur de sécurité. En 1961, Mantel et Bryan proposent pour les produits cancérigènes un modèle d'extrapolation plus complexe, basé sur la distribution semi-logarithmique: ce modèle était toutefois également applicable (en fait, il convenait mieux) aux toxiques à seuil. Le modèle de Mantel-Bryan suggérait d'effectuer une extrapolation vers les faibles doses pour calculer la dose entraînant un risque-repère de un sur 100 millions («sécurité théorique»), au moyen d'un modèle log-probit avec une pente peu prononcée, arbitrairement fixée à un écart-type par unité de log-dose (Mantel, 1975). Même si ce modèle a été adopté par certains organismes de réglementation pour les produits cancérigènes sans seuil au début des années soixante-dix, il n'était pas généralement accepté pour le calcul de la DJA de toxiques systémiques à seuil.

L'étude de la NAS *Drinking Water and Health* (NAS, 1977) a également repris l'examen de l'utilisation de modèles log-probit et log-logistiques pour la détermination des limites d'exposition aux contaminants agissant comme toxiques à seuil présents dans l'eau potable. Même si les modèles pouvaient servir à estimer le risque d'effet toxique, ils exigeaient des données tirées d'études de consommation alimentaire effectuées sur un grand nombre d'animaux pendant toute leur vie, afin de définir clairement la courbe de la distribution normale des effets d'un produit toxique. Depuis, il a rarement été possible de recueillir une quantité suffisante de données à cette fin: c'est pourquoi la NAS recommande encore la méthode DSEO-FS pour la détermination de la DJA, sauf dans les cas où les données dose-effet sont disponibles en nombre suffisant (Cotruvo, 1988).

2. Méthode d'extrapolation linéaire de Gaylor

En 1983, Gaylor propose une des premières méthodologies mathématiques pour les facteurs de sécurité (1983). La prémisse de base de la méthode de Gaylor était que le facteur de sécurité représente en partie un facteur d'incertitude qui dépend de la taille de l'échantillon de l'expérience dose-effet. Le facteur de sécurité usuel de 100 peut être soit trop grand, soit trop petit, selon la taille de l'échantillon. Un échantillon de petite taille produira une limite supérieure de confiance élevée (LSC) pour la DSEO estimée, alors qu'un échantillon de grande taille produira une LSC faible pour la même DSEO. Par conséquent, la valeur du facteur de sécurité (incertitude) devrait être directement proportionnelle à la LSC de la DSEO et inversement proportionnelle au niveau de risque tolérable prévu. Le risque peut être prédéfini à 10^{-4}, par exemple. La méthode de Gaylor propose de calculer le facteur de sécurité de la façon suivante:

$$F = LSC/R,$$

où *F* représente le facteur de sécurité; LSC, la limite supérieure de confiance à 95% pour l'effet à une dose donnée; et *R*, le niveau de risque spécifié.

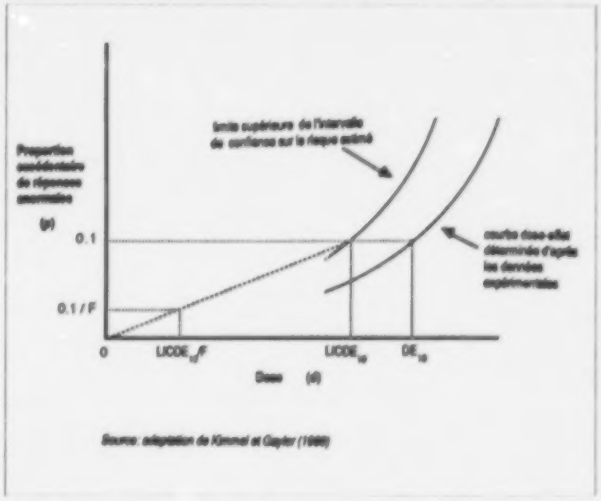
Dans bon nombre de cas, la DSEO se situe très loin vers l'extrémité inférieure de la courbe dose-effet et ne peut pas être estimée avec précision par l'observation directe. Au lieu, la méthode de Gaylor choisit, comme niveau artificiel de dose de référence ou *«pseudo-DSEO»*, la DE_{01}, qui est la dose entraînant un taux d'effet toxique de 1%. Même si la DE_{01} n'est pas souvent observable à cause de la faible différence qui existe entre ce point et le taux d'effet de base, la LSC de la DE_{01} peut être aisément calculée à l'aide des données dose-effet par un modèle mathématique d'ajustement de courbes. Tout modèle traditionnel d'extrapolation vers les faibles doses peut servir, à condition que le modèle approprié soit choisi de façon que la courbe observée suive bien la courbe théorique, ou de façon à correspondre à un mécanisme biologique plausible. Comme la courbe dose-effet est souvent concave vers le haut à faibles doses, la LSC de la DE_{01} fournit généralement une estimation plus précise de la valeur de F que la LSC de la DSEO.

La valeur de la LSC peut être calculée sur la courbe dose-effet ajustée comme la LSC à 95% pour le risque estimé à la DE_{01}, soit l'équivalent d'une DMEO hypothétique. Il suffira ensuite de diviser la valeur de LSC par le niveau de risque prévu (par exemple, 10^{-4}) pour obtenir le facteur de sécurité *F*. Une droite peut alors être tracée du point de la courbe LSC correspondant au taux d'effet de 1% jusqu'à l'origine du graphe. À partir de cette extrapolation linéaire de la courbe dose-effet, il est possible de calculer une DJA en divisant la DE01 par la valeur de *F*:

$$DJA = DE_{01}/F$$

Une version plus récente de la méthode de Gaylor recommande cependant l'utilisation de la DE_{10}, la dose efficace correspondant à un excès de risque de 10%, comme niveau de dose de référence au lieu de la DE_{01} (Kimmel et Gaylor, 1988). Il est préférable d'utiliser la DE_{01} parce que son estimation ne requiert pas d'extrapolation de la courbe déterminée bien en-deçà des données dose-effet observables. Pour déterminer la limite inférieure de confiance pour la dose correspondant à la DE_{10} ($LICDE_{10}$), il faut obtenir la courbe correspondant à la LSC du risque estimé. La $LICDE_{10}$ représente par conséquent la dose à laquelle un risque-repère de 0,1 coupe la courbe pour la LSC du risque (Figure 3.2). Il est alors possible de calculer une dose-repère pour chaque niveau de risque prédéfini en divisant la $LICDE_{10}$ par un facteur de sécurité F choisi pour la $LICDE_{10}$. Par exemple, pour obtenir un niveau de risque de 10^{-4}, un facteur de sécurité de 1000 doit être appliqué à la $LICDE_{10}$ (c'est-à-dire $0,1/1000 = 10^{-4}$). Puisque le risque est directement proportionnel à la dose dans la région de la courbe inférieure à la $LICDE_{10}$, la dose-repère pour tout faible niveau de risque spécifié peut être calculée à l'aide de $LICDE_{10}/F$ et utilisée en guise de DJA. Cette méthode est très similaire à la technique d'interpolation linéaire déjà mise au point par Gaylor pour l'analyse des données dose-effet pour la carcinogénèse (Gaylor et Kodell, 1980).

Figure 3.2
La méthode d'extrapolation linéaire utilisant la limite supérieure de l'intervalle de confiance de la DE_{10}



Source: adaptation de Kimmel et Gaylor (1988)

En considérant la situation sous un tout autre angle, Gaylor (1988) a préconisé l'adoption du modèle multistade de la carcinogénèse comme base pour un modèle d'extrapolation vers les faibles doses pour les toxiques systémiques, d'une manière analogue à celle qui est utilisée pour les produits cancérigènes sans seuil. L'avantage principal de ce modèle d'extrapolation est l'allure généralement sigmoïdale de la courbe; cette dernière se révèle ainsi utile pour la description de la dose-effet d'un grand nombre de produits non cancérigènes, dont la courbe dose-effet est généralement concave vers le haut à faibles doses. Le modèle multistade peut être défini par une équation polynomiale décrivant la probabilité d'effet P(d) en fonction de la dose; cette équation devient linéaire à faibles doses. Au moyen de la courbe ajustée, il est possible d'estimer une dose de référence à partir de la limite inférieure de confiance relative à la dose censée entraîner une incidence de maladie de 1% ou de 10% (la $LICDE_{01}$ ou la

$LICDE_{10}$) et une dose-repère pour un niveau de risque prédéfini peut être calculée en appliquant un facteur de sécurité à la dose de référence tel que décrit plus haut. Le cas échéant, les renseignements pharmacocinétiques peuvent également être incorporés au modèle: il est alors possible de connaître quelle sera la dose effice l'organe cible, de façon à réduire l'importance du risque estimé.

L'avantage principal de la méthode de Gaylor est que l'estimation du risque à faibles doses est moins subjective qu'avec les méthodes DSEO-FS. Les avantages spécifiques de la méthode de Gaylor sont les suivants: 1) elle permet une détermination flexible du facteur de sécurité; 2) elle tient compte de l'allure et de la pente de la courbe dose-effet; 3) elle permet d'obtenir une extrapolation à faibles doses prudente et ainsi de produire des valeurs semblables à celles obtenues par la méthode du facteur de sécurité; 4) elle n'exige aucune hypothèse quant à l'allure réelle de la courbe dose-effet à faibles doses; et 5) l'importance relative de la valeur du risque calculée à partir de la méthode DSEO-FS et de celle calculée à l'aide de la méthode mathématique d'extrapolation constitue une validation indépendante de la DJA dérivée des calculs DSEO-FS.

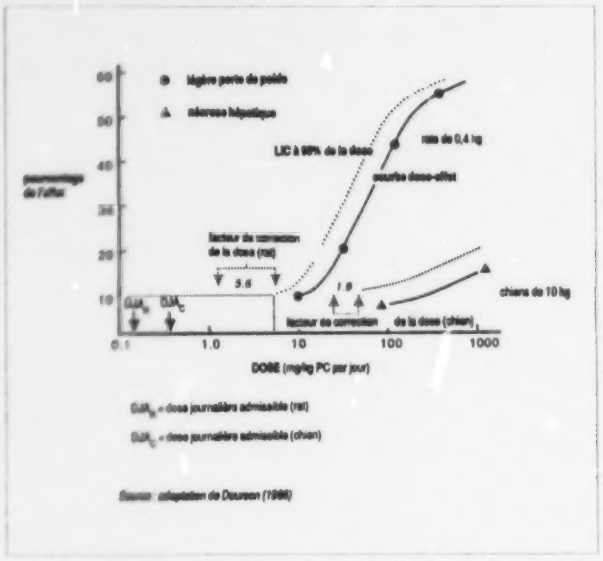
Le désavantage principal de la méthode de Gaylor est qu'elle ne s'applique qu'aux données tout ou rien et qu'elle nécessite plus de points que d'habitude pour définir convenablement l'allure de la courbe dose-effet pour la technique d'ajustement de courbe. De plus, elle ne s'applique qu'au modèle expérimental de l'animal, de sorte qu'il faut introduire un facteur de sécurité supplémentaire pour tenir compte de l'extrapolation inter-espèces de l'animal à l'homme. Un autre problème est de trouver le moyen de déterminer le niveau de risque admissible (10^{-4}, 10^{-6}) pour obtenir une dose-repère, et de trouver la façon de faire correspondre cette valeur à la notion de seuil.

3. Méthode du facteur de sécurité de Dourson

Dourson (1986) a mis au point une autre méthode pour déterminer des facteurs de sécurité flexibles. Cette méthode ressemble à celle de Gaylor au point de vue conceptuel, mais les détails diffèrent de plusieurs façons. La technique de Dourson consiste également à choisir la DE_{10} comme niveau de dose de référence: néanmoins, tout niveau de dose de référence raisonnable est également acceptable. Pour le calcul de la DJA, la courbe dose-effet est ajustée au moyen d'un modèle mathématique au choix, et la LIC à 95% de la dose correspondant au niveau de risque DE10 est déterminée d'une manière similaire à celle préconisée par Gaylor (Figure 3.3).

La DJA est alors le quotient de la LIC de la DE_{10} par plusieurs facteurs de sécurité. D'abord, un facteur de correction de la dose est introduit pour tenir compte de l'extrapolation inter-espèces: ce facteur est généralement basé sur un facteur de conversion racine cubique de l'aire. Ensuite, un facteur de 10 sert à tenir compte de la variation humaine; un second facteur de sécurité, qui peut prendre une valeur de un à 10, sert de correction pour l'utilisation d'une valeur de DMEO. La présence de ce dernier facteur s'explique par le fait que Dourson considère la DE_{10} comme l'équivalent d'une DMEO expérimentale. La valeur précise du deuxième facteur de sécurité est déterminée par la pente de la courbe dose-effet et par la gravité de l'effet toxique. Dans le cas d'une pente prononcée et d'un effet toxique grave, le FS serait de 10. Le facteur de sécurité global se situerait alors entre 10 et 100.

Figure 3.3
Méthode du facteur de sécurité de Dourson pour la détermination de la DJA



La méthode de Dourson constitue en fait une extension de la technique traditionnelle DSEO-FS mise au point par l'EPA: le point unique DSEO est remplacé par une dose de référence, représentée par la LIC à 95% de la DE_{10} dérivée d'un modèle d'ajustement de courbe. L'avantage principal de cette méthode est d'éviter de dépendre d'un seul point DSEO: elle se sert plutôt de tous les points expérimentaux pour décrire la courbe dose-effet de même que les marges d'erreur. Il est également tenu compte de facteurs tels que la pente et la gravité de l'effet. Toutefois, contrairement à la méthode de Gaylor, celle de Dourson souscrit à la méthode traditionnelle du facteur de sécurité pour le calcul de la DJA, et ne met pas en jeu d'extrapolation linéaire pour l'estimation du risque à faibles doses. De cette façon, une seule DJA peut être estimée, et il est impossible de calculer des doses-repères pour des niveaux de risque prédéfinis.

4. Méthode de la dose-repère de Crump

Crump (1984) a proposé une méthode de détermination de la DJA qui est basée sur l'analogie avec des modèles dose-effet existants pour les produits cancérigènes. Ce modèle possède un principe opérationnel semblable à celui de la méthode préconisée par Gaylor (1988). La DJA est calculée en termes de dose-repère, au lieu de la traditionnelle DSEO-FS. La dose-repère désigne *«la limite inférieure de l'intervalle de confiance statistique pour la dose correspondant à une légère augmentation de l'effet au-dessus du niveau de fond»* (Crump 1984). Le point correspondant à l'effet «légèrement au-dessus du niveau de fond» et servant à définir la dose-repère serait normalement situé près de l'extrémité inférieure de la courbe

dose-effet observable; l'augmentation de l'effet devrait être suffisamment petite pour refléter l'allure de la courbe dose-effet, mais assez grande pour éviter que la limite inférieure de confiance ne dépende trop du modèle mathématique choisi pour ce calcul. Ainsi, un niveau d'effet de un à 10% au-dessus du niveau de fond satisferait à ces exigences. Normalement, 10% constituerait le choix habituel, de sorte que la dose-repère soit l'équivalent fonctionnel de la $LICDE_{10}$ de la méthode de Gaylor (1988). La méthode de Crump permet l'application de plusieurs modèles différents d'ajustement de courbe. Il est possible de traiter à la fois des données tout ou rien, lorsqu'il ne s'agit que de la présence ou de l'absence d'un effet, et des données continues, pour lesquelles il y a une mesure de la gravité de l'effet biologique.

Dans le cas des données dose-effet *tout ou rien*, Crump suggère l'utilisation d'un modèle de *régression linéaire tout ou rien*. La probabilité d'effet $P(d)$ varie en fonction de la dose d, selon l'expression:

$$P(d) = c + (1-c)\{1 - \exp[-q_1(d-d_o)]\}$$

où c représente la constante pour le taux de l'effet de fond à un point critique, d_0 la dose seuil, et q_1 le paramètre de la pente pour la fonction dose-effet du premier degré. S'il n'est pas tenu compte de l'effet de fond, l'équation devient:

$$P(d) = 1 - \exp[-q_1(d-d_o)],$$

qui constitue une forme légèrement modifiée du modèle linéaire (1 événement délétère) utilisé pour l'extrapolation aux faibles doses des données de carcinogénèse. La variable de dose est toutefois redéfinie comme la dose située immédiatement au-dessus de la dose seuil. Par conséquent, la situation $d < d_0$ *définit la gamme de doses situées sous le seuil où le taux d'effet supérieur au niveau de fond est nul.*

D'une façon similaire, il est possible de définir, sans tenir compte du niveau de base, un modèle de *régression polynomiale tout ou rien* à l'aide d'une équation polynomiale de degré k de la forme:

$$P(d) = 1 - \exp[-q_1(d-d_o) - q_2(d-d_o)^2 - \ldots - q_k(dd_o)^k],$$

qui est une expression habituellement utilisée pour décrire le modèle multistade de carcinogénèse, et modifiée par l'introduction d'un paramètre relié à la dose seuil *do*. Une telle équation permet d'effectuer un tracé curvilinéaire à partir de données dose-effet, et confère par conséquent plus de souplesse que le modèle linéaire. Le modèle de Crump offre aussi des équations qui définissent des modèles décrivant la distribution semi-logarithmique et la distribution de Weibull.

Pour les données dose-effet *continues*, le modèle de Crump suggère des modèles d'ajustement de courbe analogues à ceux utilisés pour les données tout ou rien. Le modèle de *régression linéaire continue* proposé prend la forme suivante:

$$m(d) = c + q_1(d-d_o)$$

où $m(d)$ représente le degré d'effet au-dessus de la valeur de fond c, exprimé comme une simple fonction linéaire de la dose excédant la dose seuil d_0. Un modèle similaire de *régression polynomiale continue* peut être exprimé comme suit:

$$m(d) = c + q_1(d-d_o) + q_2(d-d_o)^2 + \ldots + q_k(d-do)^k.$$

Lorsque, pour un modèle de régression choisi, la courbe dose-effet qui le décrit et les limites statistiques d'intervalle de confiance qui le caractérisent sont fixées, il est facile de calculer la dose-repère. Celle-ci remplace la DSEO traditionnelle et désigne la limite inférieure de confiance d'une dose correspondant à une augmentation donnée du niveau de risque supplémentaire au-dessus de la valeur de fond dans le cas des données tout ou rien, ou à l'importance de l'effet supplémentaire dans le cas des données continues.

Dans le cas des données *tout ou rien*, la dose-repère désigne la dose d qui entraîne un niveau donné de risque *supplémentaire*, selon:

$$[\,P(d) - P(0)\,] \;/\; [\,1 - P(0)\,].$$

Le risque supplémentaire peut être interpété comme la probabilité d'occurrence d'un effet à une dose d, supérieur au taux d'effet de base $P(0)$ en l'absence du produit chimique étudié.

Dans le cas des données *continues*, la dose-repère désigne la dose d qui entraîne un niveau donné *d'effet supplémentaire*, selon:

$$[\,m(d) - m(0)\,] \;/\; m(0),$$

qui représente la quantité de changement dans l'importance d'un effet à une dose d par rapport au niveau de fond de l'effet.

Une fois la dose-repère déterminée, il faut introduire un facteur de sécurité pour abaisser le risque de référence du niveau 1% ou 10% à un niveau de risque «admissible». Par exemple, si la dose-repère a été fixée au niveau prédéfini de risque de 1%, il faut utiliser un facteur de sécurité de 100 pour réduire la dose-repère à un niveau de risque de 10^{-4}. Cependant, Crump ne spécifie pas la nature ou l'importance des facteurs de sécurité pour la dose-repère. Il suggère plutôt de déterminer les facteurs de sécurité de manière empirique, par comparaison des valeurs de DJA obtenues par la méthode DSEO-FS avec les valeurs de dose-repère pour les mêmes produits chimiques. Avec un nombre suffisant de comparaisons, il est possible de déterminer le facteur de sécurité approprié pour la dose-repère en calculant la moyenne des rapports dose-repère/DJA. Par la suite, ce facteur de sécurité pourrait s'appliquer à toutes les doses-repères, sauf dans les cas spécifiques exigeant le recours à des facteurs de sécurité spéciaux.

Autrement, les modèles d'ajustement de courbe peuvent servir à l'extrapolation de la courbe dose-effet vers les zones des faibles doses, pour fournir une estimation directe de la dose seuil; de même, les doses-repères correspondant à des niveaux spécifiés de risque peuvent être estimées au moyen de la courbe extrapolée. Pour établir des normes réglementaires, il est préférable d'utiliser la méthode décrite ci-haut pour le calcul des rapports coûts-bénéfices pour l'élaboration d'autres situations d'exposition, ce qui est impossible avec la méthode

traditionnelle DSEO-FS. L'incertitude qui règne cependant quant à l'allure de la courbe dose-effet à faibles doses et la sensibilité de la courbe extrapolée selon le choix du modèle ne favorisent pas cette méthode. Les estimations du seuil auront tendance à présenter des différences marquées selon le modèle choisi; ces différences sont d'ailleurs analogues à celles qui caractérisent les valeurs estimées de la *dose virtuellement sûre* (DVS) pour les produits cancérigènes.

Les avantages de la méthode facteur de sécurité-dose-repère proposée par Crump sont semblables à ceux des autres méthodes d'ajustement de courbe. Toutes les données nécessaires à la définition de la courbe dose-effet, de même que l'allure de la courbe dose-effet et ses limites de confiance statistiques servent directement à l'estimation de la dose-repère et à la dérivation de la DJA. La méthode FS-dose-repère jouit également d'une plus grande flexibilité que les autres méthodes mathématiques, parce qu'elle peut servir à l'analyse de données tout ou rien et continues, et qu'elle offre une variété de modèles d'ajustement de courbe.

5. Résumé et critique

Gaylor, Dourson et Crump ont proposé des méthodes mathématiques variées pour estimer la DJA: ces méthodes ne sont toutefois pas encore acceptées partout pour l'établissement des normes réglementaires (Dourson et Stara, 1983). Ces méthodes servent cependant pour la comparaison et la validation des valeurs de DJA dérivées de la méthode traditionnelle DSEO-FS. Malheureusement, ces modèles d'ajustement de courbe n'ont souvent pas assez de points pour assurer une estimation fiable des paramètres critiques dose-effet (par exemple, la pente) dans un grand nombre d'études aiguës ou chroniques menées sur des animaux (Dourson et coll., 1986). Dans de tels cas, les organismes de réglementation doivent continuer de compter sur la méthode traditionnelle DSEO-FS; sinon, il serait impossible de réglementer l'utilisation d'un grand nombre de produits toxiques d'usage courant dans le commerce. Crump (1986) soutient toutefois que toute valeur de DJA dérivée d'une expérience ne comportant pas suffisamment de points pour être traitée par des méthodes statistiques fondamentales devrait être considérée avec prudence. Il suggère l'utilisation du modèle dose-repère et d'autres techniques d'ajustement de courbe pour favoriser une meilleure planification des expériences et une analyse plus complète des données toxicologiques, et pour promouvoir l'utilisation plus répandue des méthodes d'évaluation quantitative du risque dans le cas des toxiques à seuil.

Par contre, les problèmes concernant le nombre restreint d'animaux et l'absence d'estimation de la pente seraient moins aigus que certains statisticiens ne le croient (Lu, 1985). Par exemple, Lu cite des études toxicologiques dans lesquelles une augmentation de 100 fois du nombre d'animaux pour une expérience dose-effet caractéristique a entraîné une réduction de 24-78% seulement, soit moins d'un ordre de grandeur, des estimations du risque-repère (le cas de la DVS pour l'éthylène thiourée) (Lu, 1985). Ces différences sont relativement insignifiantes en comparaison de la variabilité biologique qui est beaucoup plus importante, dans les études toxicologiques classiques. Lu soutient également qu'une partie du facteur de sécurité pour l'extrapolation inter-espèces est censée tenir compte de l'incertitude dose-effet provenant des différences de taille entre les groupes.

Lu recommande plutôt d'allouer plus de ressources pour la découverte d'effets pathologiques encore jamais détectés, comme par exemple la neurotoxicité retardée ou les effets tératogènes. Il conclut:

> *Il est discutable qu'une augmentation du nombre d'animaux utilisés dans les études à long terme pour accroître la taille de l'échantillon et la possibilité de déterminer la pente de la courbe dose-effet, puisse pallier le manque de tests de ce genre et d'autres modèles appropriés.*
>
> Lu (1988)

En conséquence, la question du choix des études les plus appropriées et des effets les plus pertinents pour la détermination de la DJA ne doit pas être moins importante que celle du choix du modèle et des techniques biostatistiques (Dourson et coll., 1986). Il peut être nécessaire de considérer séparément chaque effet pour le choix d'un modèle convenable, de façon que les valeurs de DJA soient fondées sur l'effet biologique le plus éloquent, en termes de gravité ou de fréquence d'occurrence. La comparaison des DJA dérivées des méthodes traditionnelles DSEO-FS et des DJA dérivées de techniques d'ajustement de courbe telles que la dose-repère peut aussi aider l'évaluateur pour le choix et la validation des facteurs de sécurité.

C. Gravité des effets

1. Échelles de gravité

Une norme d'exposition à un produit toxique peut s'obtenir d'une manière réaliste: il suffit d'accorder plus d'importance à des effets graves ou irréversibles qu'à des effets bénins ou réversibles. Par conséquent, les effets les plus graves doivent d'abord être évalués avant les effets de moindre gravité. De plus, l'importance des facteurs de sécurité pour une norme d'exposition est souvent plus rigoureuse pour les substances qui causent des effets graves que pour celles qui causent des effets bénins. Auparavant, la gravité des effets biologiques jouait un rôle significatif, quoique bien mal défini, dans la détermination des valeurs de la DSEO et des facteurs de sécurité: elle a ainsi influé sur la détermination des normes d'exposition aux DJA. Cependant, la méthodologie traditionnelle DSEO-FS ne comporte pas encore de critères ou lignes directrices normalisés pour l'inclusion de mesures de gravité dans la méthode officielle de détermination de la DJA. Les EPA Guidelines de 1988 ont différé toutes recommandations spécifiques quant à la gravité jusqu'à ce que ces méthodes soient mieux validées (Barnes et Dourson, 1988).

Tableau 3.1
Échelle de cotation d'effets utilisée pour déterminer les quantités à déclarer selon la toxicité chronique (EPA)

Échelon	Effet
1	Induction enzymatique ou autre modification biochimique sans changements pathologiques ni changement du poids des organes.
2	Induction enzymatique et prolifération subcellulaire ou autre changement au niveau des organites, mais sans changement apparent.
3	Hyperplasie, hypertrophie, ou atrophie, mais sans changement du poids des organes.
4	Hyperplasie, hypertrophie, ou atrophie, avec changement du poids des organes.
5	Modifications cellulaires réversibles: gonflement trouble, dégénérescence hydropique, changements graisseux.
6	Nécrose ou métaplasie, sans changement apparents au niveau du comportement, des sens ou de l'activité physiologique.
7	Nécrose, atrophie, hypertrophie, ou métaplasie accompagnée de diminution détectable des fonctions organiques. Toute neuropathie accompagnée de changements mesurables au niveau du comportement, des sens ou de l'activité physiologique.
8	Nécrose, atrophie, hypertrophie, ou métaplasie accompagnée de perturbation organique manifeste. Toute neuropathie accompagnée de changements évidents au niveau du comportement, des sens ou de la motricité. Toute preuve de toxicité pour le foetus.
9	Changements pathologiques prononcés accompagnés de perturbations organiques graves. Toute neuropathie avec perte du contrôle de la motricité ou du comportement, ou des facultés sensorielles. Problèmes de reproduction. Tout effet tératogène accompagné de signes de toxicité chez la mère.
10	Décès ou diminution marquée de l'espérance de vie. Tout effet tératogène sans signes de toxicité chez la mère.

Source: Hartung et Durkin (1986).

Comme l'énumération des effets n'assure qu'une description qualitative de la toxicité, il n'existe pas de méthode quantitative évidente pour évaluer la gravité d'un effet sur une échelle objective. Néanmoins, l'expérience clinique et un consensus scientifique ont permis la mise au point de plusieurs systèmes d'estimation de la gravité pour échelonner les effets de toxicité d'une manière semi-quantitative. L'EPA a élaboré une échelle de gravité basée sur la toxicité chronique: cette échelle a servi à la détermination des quantités à déclarer des produits chimiques dans les lieux de décharge de déchets toxiques (Hartung et Durkin, 1986; Stara et coll., 1987). L'échelle de *cotation des effets* mise au point par l'EPA vise principalement à l'évaluation objective de la gravité, mais elle comporte aussi certaines composantes subjectives touchant les conséquences pour la santé (Tableau 3.1). L'échelle présente une progression graduée de un (bénin) à 10 (grave) des effets délétères pour la santé. Les six premiers échelons se rapportent principalement à des modifications cellulaires sub-cliniques qui n'entraînent aucun trouble visible de la fonction physiologique. Les échelons 7, 8 et 9 représentent trois classes de trouble organique en ordre croissant de gravité: détectable, manifeste et grave. L'échelon 10 est réservé aux effets mortels. Certains effets tératogènes ou neurologiques variés font partie d'échelons particuliers en raison de l'inquiétude marquée qu'ils soulèvent dans la société.

2. Estimation de DSEO au moyen de diagrammes dose-gravité

Pour évaluer des schémas mixtes dose/gravité dans le cas des produits de toxicité générale et pour aider à l'estimation de la DJA, l'EPA a récemment imaginé de combiner la *cotation des effets* (Cot_e) et un système distinct d'évaluation du niveau de dose (puissance) d'un produit chimique donné (Stara et coll., 1987). L'échelle de cotation des effets est divisée en plusieurs niveaux d'effets dans l'ordre de gravité croissante:

- DSEDO
  dose qui n'entraîne pas d'effets délétères
  Cot_e = 1 à 3 ou 4
- Dose à effets délétères (ou DMEDO)
  dose qui entraîne des effets délétères
  Cot_e = 5 à 7
- DEP
  niveau qui entraîne des effets patents
  Cot_e = 8 à 10

Ces effets sont expliqués plus en détail dans l'échelle (EPA) descriptive des effets toxiques (Tableau 3.1) élaborée par Hartung et Durkin (1986).

Ensuite, une *Cotation des doses* (Cot_d) est définie en fonction de la *dose minimale efficace* (DME). La DME constitue un indice général de la puissance toxique d'un produit chimique donné en termes de dose (exprimée en milligrammes par jour) nécessaire pour causer un effet détectable d'un niveau donné de gravité. La DME est convertie en valeurs comprises entre 1 et 10 sur une échelle logarithmique, selon l'expression:

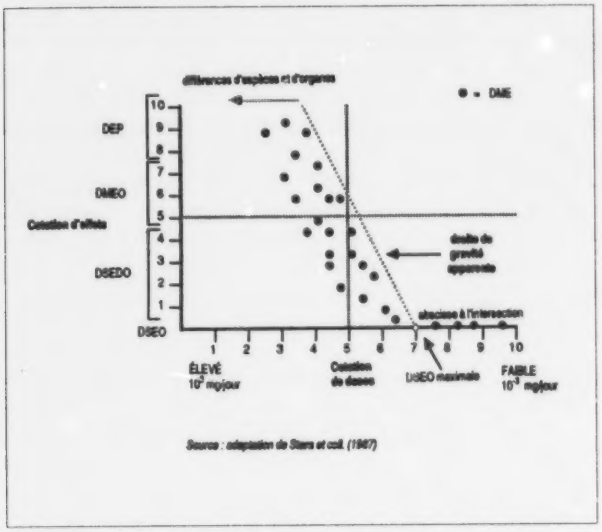
$$RV_d = -1{,}5 \log DME + 5{,}5$$

qui couvre les valeurs de DME comprises entre -3 et 3. À l'extérieur de cet intervalle, la Cot_d est fixée à 1 pour une DME supérieure à 10^{-3} mg/j, et elle est fixée à 10 pour une DME inférieure à 10^{-3} mg/j. Comme la puissance est inversement proportionnelle à la DME, la substance qui exerce un effet à faible dose aura une cote élevée sur l'échelle Cot_d.

Il devient alors possible de représenter graphiquement la relation dose-effet pour un produit chimique sur un diagramme de dispersion bidimensionnel des valeurs de Cot_e et de Cot_d correspondant à divers effets biologiques au niveau de différents organes (Figure 3.4). Chaque point expérimental du diagramme des Cot_d/Cot_e représente la valeur de la DME pour un effet biologique particulier au niveau d'un organe spécifique ou pour une espèce donnée. Lorsque la dose augmente

(c'est-à-dire la Cot_d diminue), la cote de l'effet augmente, depuis les effets bénins correspondants à la DSEDO jusqu'aux effets modérés correspondant à la DMEDO et finalement aux effets graves correspondant à la DEP. Les points du diagramme représentent une variété d'effets biologiques différents: l'augmentation de la dose favorise la manifestation d'effets multiples.

Figure 3.4
Estimation de la DSEO par diagrammes dose-gravité construits au moyen du tracé de la Cotation d'effets en fonction de la Cotation de doses



Pour un niveau Cot_e donné, le point le plus éloigné vers la *droite* représente l'occurrence d'un effet au niveau de l'organe le plus sensible ou chez l'espèce la plus sensible. Par conséquent, la ligne tracée à droite de tous les points expérimentaux de Cot_d/Cot_e définit la dose minimale requise pour causer un effet de gravité variable au niveau de l'organe le plus sensible ou chez l'espèce la plus sensible. La pente de cette droite, appelée *droite de gravité* apparente, peut être extrapolée vers le bas pour croiser l'axe des × juste à la gauche de plusieurs valeurs de DSEO pour différents organes et espèces. L'abscisse à l'intersection de cette droite représente donc la DSEO maximale, qui est alors considérée comme la DSEO du produit chimique étudié, et qui tient compte de l'effet produit à tous les niveaux de gravité et pour de multiples effets biologiques au niveau de différents organes et chez diverses espèces.

La méthode du tracé de la Cot_d/Cot_e offre plusieurs avantages par rapport aux tracés traditionnels dose-effet pour l'obtention des DJA. Il est possible de traiter le tracé de la Cot_d/Cot_e en tant que relation quantitative dose-effet, parce que les valeurs de Cot_e et de Cot_d sont destinées à être des approximations d'une échelle d'intervalles, plutôt que seulement une échelle graduée de la dose et la gravité de l'effet. La détermination de la DSEO et de la DJA est alors rendue possible par l'application de méthodes statistiques paramétriques pour estimer la droite de gravité apparente et l'abscisse à l'intersection, et en même temps toutes les données provenant de plusieurs expériences chroniques et subchroniques pour différents effets à des niveaux variés de gravité. De plus, cette méthode s'avère particulièrement utile en l'absence de valeurs adéquates de DSEDO ou de DMEDO tirées des données expérimentales. Le cas échéant, les données basées sur les DEP et les DSEO peuvent être utilisées pour estimer la DSEO maximale ou une DSEDO appropriée (une dose associée à une Cot_e d'une valeur de un ou deux).

La méthode décrite plus haut a été élaborée par le personnel scientifique de l'EPA dans le but de tenir compte de la gravité de l'effet dans l'estimation de la DSEO. Même si elle ne fait pas encore officiellement partie des lignes directrices de l'EPA pour la détermination par DSEO-FS des DJA, la méthode a été appliquée à l'évaluation de la relation dose/gravité pour plus de 250 toxiques systémiques; dans l'avenir, elle pourrait même constituer une méthode de rechange efficace aux méthodes classiques de détermination de la DJA.

4. Méthodes FS appliquées aux produits cancérigènes

A. Généralités

Au cours des deux dernières décennies, les méthodes classiques d'évaluation de risque pour les produits cancérigènes étaient fondées sur l'hypothèse que tous les produits cancérigènes sont génotoxiques et qu'ils agissent selon un processus stochastique en endommageant l'ADN cellulaire au hasard. Par conséquent, la plupart des modèles d'évaluation du risque présenté par un produit cancérigène sont conformes à l'hypothèse d'une linéarité sans seuil, selon laquelle il n'existe pas de seuil pour l'induction du cancer, même pour des expositions aux doses les plus faibles et selon laquelle la probabilité d'induction de cancer à de faibles niveaux d'exposition à un produit cancérigène est directement proportionnelle à la dose. Pour cette raison, les méthodes d'évaluation quantitative du risque (EQR) pour les produits chimiques cancérigènes ont utilisé des modèles mathématiques d'extrapolation vers les faibles doses pour l'estimation du risque cancérigène, fondés sur la notion de la dose virtuellement sûre (DVS), une dose-repère correspondant à un risque «admissible» de cancer au cours d'une vie, de 10^{-6} ou de 10^{-5}.

La méthodologie EQR ne comprend aucun facteur de sécurité explicite. Mais les normes d'exposition calculées sont extrêmement rigoureuses, en raison du choix des modèles d'extrapolation les plus prudents (linéaire 1-événement délétère ou multistade linéarisé) et de l'utilisation des limites statistiques supérieures (limite de confiance à 95%) de la courbe dose-effet pour l'estimation du risque de carcinogénicité. La prudence inhérente aux hypothèses comprises dans les modèles EQR peut être considérée comme l'équivalent des facteurs de sécurité utilisés dans la méthodologie DSEO-FS.

Un grand nombre d'organismes de réglementation, et particulièrement la U.S. EPA qui s'était faite le défenseur des méthodes EQR pour réglementer l'utilisation des produits cancérigènes, ont rejeté les méthodes DSEO-FS pour la détermination des limites d'exposition dans le cas des produits cancérigènes (Anderson, 1985; EPA, 1986b). De la même façon, la plupart des lignes directrices établies par divers organismes consultatifs aux États-Unis ont toujours appuyé l'hypothèse de linéarité sans seuil et son application à l'évaluation du risque de carcinogénicité (IRLG, 1979; OSTP, 1986). Dernièrement, l'EPA et d'autres organismes américains de réglementation ont toutefois réexaminé l'utilisation possible des méthodes DSEO-FS dans des cas spéciaux pour lesquels l'hypothèse de la linéarité sans seuil ne serait pas pertinente.

En pratique, l'EPA a récemment proposé le recours à des normes réglementaires révisées concernant d'importantes substances cancérigènes, pour des motifs reliés aux mécanismes ou à la pharmacocinétique : il est alors devenu possible d'utiliser des modèles à seuil comprenant un facteur de sécurité. Dans certains cas, d'autres organismes de réglementation au Canada et ailleurs ont fait de même.

B. Facteurs traditionnels de sécurité pour produits cancérigènes

À l'occasion, les méthodes traditionnelles DSEO-FS ont servi à la détermination de limites d'exposition pour les produits cancérigènes, avec l'inclusion de facteurs de sécurité tenant compte de l'incertitude additionnelle générée par les propriétés exceptionnelles des substances cancérigènes. Weil (1972) a d'abord proposé un facteur de sécurité de 5000 pour ces produits. Cette valeur comprenait le facteur de sécurité de 100 habituel, un facteur de 5 qui tenait compte du fait que la plus faible dose administrée représentait une DMEO plutôt qu'une DSEO, et enfin un facteur de 10 qui reflétait l'irréversibilité générale du processus de la carcinogénèse.

Le ministère de la Santé nationale et du Bien-être social du Canada a déterminé, à plusieurs reprises, des limites réglementaires d'exposition à des produits cancérigènes, à l'aide des méthodes DSEO-FS appliquées cas par cas (Somers, 1981, 1984). En 1981, la limite permise de concentration de TCDD (dioxine) présente dans le poisson commercial était fixée à 20 ppb: le calcul était basé sur un facteur de sécurité de 2000 appliqué à la valeur de DMEO obtenue d'expériences dose-effet menées sur des rongeurs. Le raisonnement qui sous-tendait l'emploi des méthodes DSEO-FS était fondé sur l'insuffisance de preuves quant à la présence d'un mécanisme génotoxique pour la dioxine au cours d'épreuves de mutagénicité, quant à l'absence de réactivité avec l'ADN et quant à son mécanisme d'action mettant en jeu l'induction

enzymatique et peut-être d'autres effets épigénétiques de régulation. De la même façon, la ligne directrice quant à la présence de formaldéhyde dans l'air ambiant à l'intérieur des habitations a été fixée à 0,1 ppm, d'après une réduction de 20 fois de la norme existante concernant l'air ambiant en milieu de travail.

Il est évident que ces décisions ne constituent pas une façon rigoureuse d'aborder le problème de l'évaluation du risque présenté par les produits cancérigènes, mais elles permettent de déterminer simplement les limites d'exposition à ces produits qui ont des comportements dose-effet à «seuil». Cette façon d'aborder le problème n'a toutefois pas été retenue, pour des raisons obscures, dans le cas de l'établissement des limites d'exposition à l'acide nitriloacétique (NTA) dans l'eau potable. Le NTA agit comme produit cancérigène au niveau du rein et de la vessie seulement chez les rats à des doses extrêmement élevées qui induisent un effet cytotoxique au niveau des voies urinaires, des animaux de laboratoire. L'effet cytotoxique a été interprété comme un précurseur de l'induction de tumeurs des voies urinaires. L'utilisation de NTA en tant qu'additif aux détergents à lessive a été permise au Canada (mais non aux États-Unis) sous prétexte que les doses tellement plus faibles de NTA habituellement présentes dans l'eau potable ne seraient pas susceptibles d'induire des effets cytotoxiques et que par conséquent l'induction de cancer était exclue. L'évaluation du risque effectuée par Santé et Bien-être social Canada a toutefois été faite à l'aide d'un modèle classique d'extrapolation vers les faibles doses, même en l'absence de preuves de la génotoxicité du NTA ou de sa réactivité avec l'ADN.

Plus récemment, l'American Conference of Government Industrial Hygienists (ACGIH) a publié un article de politiques décrivant les lignes directrices révisées de l'ACGIH pour les produits cancérigènes (ACGIH, 1986), et sa tradition de longue date qui consiste à appliquer les méthodes FS pour l'évaluation du risque présenté par ces produits. Selon ces lignes directrices, la méthode de classification des produits cancérigènes préconisée par l'ACGIH groupe ces substances en deux catégories:

A1: *produits cancérigènes confirmés pour l'homme* selon des preuves épidémiologiques chez l'homme

A2: *produits cancérigènes suspectés pour l'homme* selon des preuves épidémiologiques partielles ou selon la carcinogénicité chez l'animal de laboratoire

Pour l'établissement des normes d'exposition en milieu de travail exprimées sous forme de valeurs limites d'exposition (VLE), l'ACGIH recommande qu'une VLE ne soit pas attribuée aux produits cancérigènes A1, en raison du manque de preuves expérimentales d'un niveau d'exposition admissible. En fait, il s'agit ici de la philosophie de tolérance nulle, par laquelle la substance est interdite ou gardée hors du contact avec le travailleur par des dispositifs de protection. Dans le cas des produits cancérigènes A2, l'ACGIH recommande le recours à une méthode DJA, dans laquelle la limite d'exposition est déterminée par le quotient de la DSEO par un facteur de sécurité arbitraire de 100 ou de 1000 fournissant une marge de sécurité convenable. Il n'est fait mention d'aucune méthode pour la spécification d'un facteur de sécurité approprié.

Le raisonnement exposé pour cette philosophie est fondé sur l'argument biologique voulant que les défenses de l'organisme puissent neutraliser des faibles doses de produits cancérigènes, et sur l'argument méthodologique voulant que les modèles mathématiques pour l'extrapolation vers les faibles doses soient trop controversés et trop subjectifs pour être utilisés pour l'établissement de normes. Le modèle classique à seuil de toxicité est alors considéré comme applicable aux produits cancérigènes, malgré le fait que la plupart des autres organismes scientifiques et ceux chargés de la réglementation ne soient pas en accord avec cette hypothèse simpliste.

C. Méthodes de classification selon l'importance des preuves

Les EPA *Guidelines for Carcinogen Risk Assessment* de 1986 (Lignes directrices relatives à l'évaluation du risque présenté par les produits cancérigènes) (EPA, 1986a) contiennent un système de classification des produits cancérigènes qui définit les critères d'attribution des divers niveaux de «potentiel de carcinogénicité» s'appliquant à des produits cancérigènes spécifiques (Tableau 4.1). Un système de classification comprenant cinq groupes (Groupe A à Groupe E) décrit le degré de certitude qu'un produit chimique donné représente un véritable produit cancérigène pour l'homme, à la lumière des données épidémiologiques et des résultats de bioessais chroniques.

Tableau 4.1
Lignes directrices de l'EPA classification des produits cancérigènes selon leur potentiel de carcinogécité

Groupe A -	*Produit cancérigène pour l'homme*: nombre suffisant de preuves expérimentales chez l'homme.
Groupe B -	*Produit probablement cancérigène pour l'homme*: nombre limité de preuves expérimentales chez l'homme ou aucune preuve expérimentale chez l'homme et nombre suffisant de preuves expérimentales chez l'animal.
Groupe C -	*Produit peut-être cancérigène pour l'homme*: nombre limité de preuves expérimentales ou preuves équivoques chez l'animal en l'absence de données relatives à l'homme.
Groupe D -	*Carcinogénicité non déterminée chez l'homme*: données inappropriées ou absentes.
Groupe E -	*Pas de preuves expérimentales de carcinogénicité pour l'homme*: preuves expérimentales négatives pour au moins deux espèces.

Source: EPA (1986b).

En plus de servir à l'identification des produits cancérigènes potentiels, le système de classification mis au point par l'EPA a aussi été utilisé pour le choix des techniques méthodologiques appropriées dans la détermination des normes relatives aux contaminants présents dans l'eau potable préconisées par l'EPA — y compris les produits organiques de synthèse, les produits organiques volatils et les produits inorganiques (EPA, 1985). Cette norme ainsi recommandée, auparavant connue sous le terme «limite maximale de contaminant recommandée» (LMCR), est maintenant appelée un «objectif de dose maximale de contaminant» (ODMC).

Le EPA Office of Drinking Water (Bureau de l'eau potable) utilise actuellement un système comportant trois catégories pour établir les ODMC (Tableau 4.2) en fonction de la classification des produits cancérigènes selon leur potentiel de carcinogénicité (Cotruvo, 1988).

Tableau 4.2
Méthode élaborée par l'EPA comprenant trois catégories et servant à obtenir les objectifs de dose maximale de contaminant (ODMC) dans le cas des contaminants présents dans l'eau potable

Catégorie I («produits cancérigènes probables»)
substances des groupes A et B
ODMC nul (objectif souhaitable)

Catégorie 2 («produits de carcinogénicité incertaine»)
substances du groupe C
ODMC égal à la DVS à un excès de risque de cancer de 10^{-5} à 10^{-6}
ou
ODMC égal à la DJA avec des facteurs de sécurité additionnels

$$DJA = DSEDO / 1000$$

FS normalisé de 100 fois et un FS additionnel de 10 fois pour les produits chimiques dont «les preuves de carcinogénicité sont incertaines»,
ou
un facteur de sécurité autre que 10 si les propriétés du produit chimique le prescrivent

Catégorie 3
substances des groupes D et E
traités comme des produits toxiques à seuil (non cancérigènes)
méthodologie de la DJA avec une portion attribuée à l'eau potable.

Source: Cotruvo (1988).

La catégorie I («produits cancérigènes probables») comprend les produits cancérigènes des groupes A et B pour lesquels les preuves sont très valables (importantes). L'ODMC de toute substance de la catégorie I est la concentration nulle, une valeur idéalisée difficile à obtenir en pratique. Ainsi, les ODMC sont considérés comme des objectifs «souhaitables», mais non comme des normes.

La catégorie II («produits de carcinogénicité incertaine») comprend les substances du groupe C dont les preuves sont de moindre valeur. L'EPA permet la détermination des ODMC pour les substances de la catégorie II de deux manières: soit par le calcul, au moyen d'un modèle prudent d'extrapolation vers les faibles doses, de la DVS qui entraîne un excès de risque de cancer de 10^{-5} ou 10^{-6} au cours d'une vie entière, soit par le calcul de la valeur de la DJA (DRéf) avec des facteurs de sécurité additionnels qui assurent une «marge de sécurité convenable». Tous les ODMC calculés par la méthode DSEO-FS se rapportent à des effets non cancérigènes. Le mode de détermination préféré est la méthode DSEO-FS en présence de données non cancérigènes adéquates. Il s'agit de calculer la DJA à l'aide du facteur de sécurité normal de 100 fois; ensuite, selon chaque cas, il est tenu compte de l'incertitude additionnelle au moyen d'un facteur de sécurité variant de 1 à 10.

Cette méthodologie a été appliquée pour la détermination des LMCR de six produits organiques de synthèse dans le *National Primary Drinking Water Regulations* de 1985 (EPA, 1985). Les six produits chimiques classés dans la catégorie II sont: le styrène, le 1,2-dichloropropane, le monochlorobenzène, le lindane, l'amiante et l'arsenic. Pour les produits de cette catégorie, la DJA résulte habituellement du quotient DSEDO par un facteur de sécurité de 1000, comprenant le facteur de sécurité de 100 et un facteur de sécurité additionnel de 10, pour les produits chimiques dont les «preuves de carcinogénicité sont incertaines» (EPA, 1985). Un facteur d'une valeur autre que 10 peut toutefois être appliqué si les propriétés du produit chimique l'exigent.

Par exemple, la LMCR du styrène provient de données de toxicité chronique parce que les résultats du bioessai de carcinogénicité étaient très limités et inappropriés pour l'extrapolation. La DJA du styrène est le résultat du quotient de la DSEDO par un facteur de sécurité de 1000, comprenant un facteur de sécurité supplémentaire de 10 pour tenir compte de l'exposition pendant une partie de la vie des animaux de laboratoire (EPA, 1985). La LMCR du styrène a ensuite été déterminée en divisant la valeur de la DJA (7 mg/L) par un facteur de sécurité additionnel de 10 tenant compte de la carcinogénicité potentielle du styrène, et par un facteur d'exposition de 5 assumant la contribution de l'eau potable à 20% de l'apport total. Le quotient de la DJA par 50 donne enfin la LMCR finale, qui est de 0,14 mg/L.

La catégorie III comprend les substances des groupes D et E, pour lesquels la probabilité d'action cancérigène est faible ou négligeable. Ces produits chimiques sont considérés comme des produits toxiques à seuil (non cancérigènes): la méthode de la DJA leur est appliquée, avec toutefois une correction apportée pour tenir compte de la proportion de l'apport total attribué à l'eau potable. Par exemple, le solvant 1,1,1-trichloroéthane (groupe D selon l'EPA) a été classé dans la catégorie III, et son utilisation a été réglementée en tant que celle d'un produit non cancérigène selon les normes de l'EPA de 1987 pour l'eau potable spécifiées pour huit produits organiques volatils (Cotruvo, 1988).

D. Système de classification mécanistique

Les méthodologies traditionnelles d'évaluation du risque présenté par un produit cancérigène reposent sur l'hypothèse que tous les produits cancérigènes agissent par le biais d'une atteinte mutagène au niveau de l'ADN cellulaire, et que tous les produits cancérigènes devraient donc être considérés comme produits génotoxiques. Au cours des dernières années, un ensemble grandissant de preuves a toutefois montré que la carcinogénèse peut avoir lieu selon plusieurs modes d'action. Il s'est alors ensuivi une classification des produits chimiques en deux catégories de base, selon leur mécanisme d'action: 1) les produits cancérigènes *génotoxiques*, qui causent des mutations somatiques soit en endommageant directement l'ADN, soit en réagissant indirectement avec d'autres composants génétiques de la cellule, et 2) les produits cancérigènes *épigénétiques*, qui affectent la régulation cellulaire et le génotype sans altérer l'ADN-même (Weisberger et Williams, 1981).

D'après une interprétation relativement simpliste du modèle multistade de la carcinogénèse, les initiateurs de tumeurs sont généralement considérés comme des produits cancérigènes génotoxiques, alors que les promoteurs de tumeurs, les hormones cancérigènes, les agents cytotoxiques et les agents favorisant la prolifération des peroxysomes sont classés comme des produits cancérigènes épigénétiques. La principale conséquence d'une telle distinction entre les mécanismes d'action est associée au comportement dose-effet du produit chimique en question. On suppose que les produits cancérigènes génotoxiques suivent une relation dose-effet linéaire sans seuil et présentent un risque non nul à faibles doses. Par contre, on croit que les produits cancérigènes épigénétiques possèdent un seuil de toxicité cellulaire, en-dessous duquel le risque d'induction de cancer est nul (Weisberger et Williams, 1981):

> *La plupart des agents épigénétiques ne sont actifs qu'à fortes doses soutenues et, dans une certaine mesure, les effets qu'ils induisent sont réversibles. Par conséquent, ces types d'agents peuvent ne représenter qu'un risque quantitatif pour l'homme: il peut être possible d'établir des niveaux sûrs d'exposition après qu'on aura effectué des études toxicologiques dose-effet appropriées.*
>
> Weisberger et Williams (1981)

Depuis que la distinction génotoxique/épigénétique a été faite au début des années quatre-vingt, plusieurs organismes de réglementation ont pensé instaurer un système de classification mécanistique pour réglementer l'utilisation des produits cancérigènes. La portée de ce système de classification pour l'évaluation du risque présenté par les produits cancérigènes sur la réglementation est considérable. En principe, le cas des produits cancérigènes épigénétiques pourrait être adapté aux méthodes DSEO-FS pour l'établissement de normes réglementaires, fournissant ainsi au moins une assurance nominale de «sécurité absolue» pour cette classe de produits. Qui plus est, les estimations du risque faites selon les méthodes DSEO-FS sont souvent inférieures de plusieurs ordres de grandeur à celles résultant de modèles prudents d'extrapolation vers les faibles doses (Perera, 1984). D'où le fait que les limites d'exposition calculées pour les produits cancérigènes épigénétiques peuvent être plusieurs milliers de fois moins rigoureuses que celles déterminées de modèles d'extrapolation vers les faibles doses. Cependant, cet avantage plus ou moins évident peut être fortement amenuisé par l'introduction de facteurs de sécurité additionnels tenant compte de l'incertitude générée lors de la détermination de la DSEO dans les études de bioessais de carcinogénicité, et reflétant l'irréversibilité de l'action cancérigène (EPA, 1985).

Les débats scientifiques et politiques les plus importants au sujet de l'application de la distinction génotoxique/épigénétique et de l'utilisation des méthodes DSEO-FS ont eu lieu aux États-Unis, et plus particulièrement au sein de l'EPA (Marshall, 1982; Perera et Petito, 1982; Ashford et coll., 1983, Perera, 1984). La EPA Cancer Policy de 1976 (Politique de l'EPA sur le cancer) proposait de s'en tenir rigoureusement au modèle linéaire sans seuil pour l'évaluation du risque présenté par tous les produits cancérigènes. En 1979, cette politique avait de plus reçu l'appui du rapport de l'Interagency Regulatory Liaison Group (IRLG, 1979). Les lignes directrices de l'IRLG concernant l'évaluation du risque présenté par les produits cancérigènes appuyaient sans réserve la position de l'EPA en affirmant qu'il était impossible de supposer que les produits cancérigènes avaient un comportement à seuil. Seule une reconnaissance prudente d'interprétations éventuelles des données expérimentales selon le mécanisme d'action était permise dans des cas exceptionnels.

Toutefois, en 1982, les administrateurs de l'EPA ont tenté de mettre de l'avant une proposition préliminaire comprenant des révisions de la politique concernant les méthodes d'évaluation du risque utilisées par l'Agence, sous la forme d'un «Addendum» à la EPA Cancer Policy de 1976 (Marshall, 1982; Perera et Petito, 1982). Cet addendum préconise un nouveau système de classification de la génotoxicité, fondé sur les résultats positifs obtenus à deux épreuves différentes de mutagénicité *in vitro* (l'une, sur des microbes, l'autre, sur des eucaryotes), ou à une seule épreuve positive effectuée sur des mammifères *in vivo*. Le concept de base de la proposition était le système de classification mécanistique élaboré par Weisberger et Williams (1981); cette proposition comportait cependant un nombre plus restreint d'épreuves sans spécification quant à une batterie d'épreuves de mutagénicité.

Au plan de la réglementation, l'application la plus importante de la nouvelle politique de l'EPA était une proposition à l'effet que le formaldéhyde (FA) gazeux devait être réglementé en tant que produit cancérigène épigénétique. Le mémorandum Todhunter de 1982 recommandait que le FA devait être réglementé par l'EPA en tant que produit cancérigène à seuil, à la lumière de la cytotoxicité qu'il présentait au niveau de l'épithélium nasal chez l'animal de laboratoire et chez l'homme (Perera et Petito, 1982). La proposition d'une norme d'exposition établie en fonction du mécanisme d'action du FA citait l'article de Weisberger et Williams (1981) comme raison d'être de la distinction

épigénétique/génotoxique. Le mémorandum de Todhunter a cependant été critiqué sévèrement pour des raisons scientifiques et pour des raisons méthodologiques à la fois par l'EPA et ses organismes consultatifs: puis le mémorandum a été retiré (Perera et Petito, 1982; Ashford et coll., 1983). La politique mécanistique du cancer de l'EPA a été abandonnée plus tard, par suite du manque de consensus scientifique et en raison de troubles administratifs au sein de l'organisme (l'affaire «Sewergate»).

Les tentatives suivantes faites dans le système de réglementation des États-Unis pour réinstaurer les méthodes mécanistiques ont été confrontées à des doutes sérieux émis par le milieu scientifique. Au moment de la première ébauche du rapport de l'Office of Science Technology Policy au sujet de la carcinogénèse chimique (OSTP, 1986), les méthodes mécanistiques ont été proposées une nouvelle fois; elle ont été toutefois reléguées aux oubliettes à cause des critiques émises par un grand nombre de scientifiques tant en réglementation qu'en recherche (Rushefsky, 1985). Des heurts similaires ont ponctué l'ébauche de l'*EPA Guidelines for Carcinogen Risk Assessment* de 1986 (EPA, 1986b).

Parmi les critiques les plus sérieuses, on trouve la condition selon laquelle un seuil toxicologique doit exister pour que les promoteurs et les autres produits cancérigènes épigénétiques puissent agir. Cette prémisse a été faiblement validée, en raison de l'absence d'études dose-effet appropriées, et parce que les études de promoteurs existantes ne sont généralement pas conçues en fonction de l'analyse dose-effet (OSTP, 1986). À un niveau plus fondamental, Perera (1984) a critiqué le concept d'une distinction épigénétique/génotoxique en produits cancérigènes en raison du fait que les promoteurs peuvent agir selon des mécanismes *génotoxiques*. Ses arguments peuvent se résumer comme suit:

- il n'y a pas de démarcation nette ou de distinction biologique rigoureuse entre les produits cancérigènes «génotoxiques» et «épigénétiques»
- il est impossible d'assumer qu'un seul mécanisme plausible ou significatif soit nécessairement génotoxique ou que les situations épigénétique/génotoxique s'excluent l'une l'autre
- il est simpliste d'associer l'activité génotoxique à l'initiation et l'activité épigénétique à la promotion
- les courbes dose-effet ne sont pas nécessairement différentes dans les cas des produits cancérigènes épigénétiques; il n'est pas possible non plus de supposer l'existence d'un seuil. En conséquence, les produits cancérigènes épigénétiques ne peuvent pas être considérés comme plus «sûrs» que les produits cancérigènes génotoxiques
- un grand nombre de promoteurs ont un pouvoir d'induction de tumeurs très élevé (par exemple, le TPA et la TCDD); ces produits peuvent se révéler aussi actifs que les produits cancérigènes classiques puissants.

De plus, le rapport de l'OSTP souligne que les seuils d'une population pour un produit cancérigène épigénétique sont variables même s'il est possible de déterminer un seuil chez l'animal de laboratoire (OSTP, 1986). Cette objection peut toutefois s'appliquer au cas de tous les toxiques classiques; un facteur de sécurité approprié pour une population peut cependant en tenir compte. Pour le moment, les arguments pour et contre les méthodes mécanistiques d'évaluation du risque peuvent être résumées comme suit:

> *Les opposants à ce point de vue soutiennent que les études négatives menées sur la batterie actuelle (d'épreuves) de génotoxicité par ces prétendus produits cancérigènes épigénétiques ne prouvent pas de façon claire que l'agent n'agit pas au niveau de l'ADN ou qu'il agit selon un seuil. Ils soutiennent que le manque de sensibilité des épreuves actuelles pour distinguer la faculté de tous les produits cancérigènes à initier l'ADN ne devrait pas constituer une excuse pour instaurer un nouveau système de convictions dont les erreurs ne vont pas précisément dans le sens de la sécurité. Les tenants de la position épigénétique répliquent en affirmant qu'il est clairement impossible de prouver la négative et que pour des raisons d'orientation philosophique, de tels scientifiques ne seraient toujours pas satisfaits d'une ou de deux douzaines de nouvelles épreuves négatives de génotoxicité.*
>
> *Cette question peut rester sans réponse, mais ses implications au niveau de l'économie et de la santé publique lui confèrent une importance de taille. Si la position épigénétique demeure, on verra émerger le problème de la définition d'un facteur de sécurité acceptable pour les produits cancérigènes.*
>
> Calabrese (1983)

Les EPA *Guidelines for Carcinogen Risk Assessment* de 1986 (EPA, 1986b) affichent une permissivité prudente au sujet des méthodes mécanistiques d'évaluation du risque, en autorisant l'utilisation de facteurs de sécurité dans des circonstances où les preuves concernant l'action du promoteur ou un autre mécanisme épigénétique sont particulièrment convaincantes:

> *En présence de données pharmacocinétiques ou métaboliques, ou de toute autre preuve substantielle de l'existence d'un mécanisme d'action de la carcinogénèse, un modèle d'extrapolation vers les faibles doses autre qu'une procédure multistade de linéarisation peut se révéler la mieux appropriée pour des raisons biologiques.*
>
> EPA (1986b)

Le recours aux méthodes FS pour l'évaluation du risque présenté par les produits cancérigènes est illustré au mieux par des révisions récentes des normes d'exposition à la dioxine (TCDD) dans l'eau potable, effectuées à la fois par l'EPA aux États-Unis et le Ministère de l'environnement de l'Ontario au

Canada. Un exemple de recours aux méthodes pharmacocinétiques est la modélisation pharmacocinétique à base physiologique (PBP) qu'utilise l'EPA pour le solvant chlorure de méthylène. Ces méthodologies seront étudiées plus en détail dans les prochaines sections.

E. Méthodes mécanistiques d'évaluation du risque appliquées au cas de la dioxine

1. Méthodologie de l'EPA

Des changements importants au niveau de la méthodologie de l'EPA relative à l'évaluation du risque présenté par les agents cancérigènes ont été proposés récemment pour réviser les limites d'exposition à la dioxine — spécifiquement la 2,3,7,8-tétrachlorodibenzo-p-dioxine, ou TCDD (EPA, 1987; Shu et coll., 1987; Gough, 1988; Finkel, 1988). À l'EPA, les estimations de risque existantes et revues sont toutes les deux établies à partir des mêmes données dose-effet de tumeur hépatique induites par la TCDD chez le rat (Kociba et coll., 1978). Dans ces études, un effet patent de tumeur hépatique a été observé à une dose de 0,1 µg/kg/jour, un effet léger de nature non tumorale se présentant sous forme d'induction enzymatique au niveau du foie a été observé à 0,01 µg/kg/jour, et aucun effet n'a été observé à 0,001 µg/kg/jour. L'application d'un modèle traditionnel d'extrapolation linéarisée multistade à ces données sur la TCDD produit une limite de dose journalière extrêmement rigoureuse de 0,006 *picogrammes* (pg)(1 pg = 10^{-6} µg) par kilogramme (kg) de poids corporel par jour (EPA, 1985).

Une telle limite restrictive d'apport pour la dioxine équivaut à peu près à une dose journalière de 10^{-7} molécules de TCDD par kilogramme de poids corporel, une quantité qui n'est susceptible d'exercer aucun effet biologique ou cancérigène significatif, à moins que la molécule ne soit un agent génotoxique puissant. En fait, des batteries complètes d'épreuves de mutagénicité *in vitro* et *in vivo* révèlent que la TCDD ne présente pratiquement pas d'activité génotoxique ou initiatrice de tumeur et qu'elle ne démontre aucune preuve d'interaction chimique avec l'ADN (Shu et coll., 1987).

À l'opposé, la TCDD s'est révélée agir comme un puissant promoteur dans plusieurs systèmes, dont le foie de rat préalablement initié par la diéthylnitrosamine, et la peau de souris initiée par des hydrocarbures aromatiques polycycliques (Gough, 1988). Ces résultats suggèrent que la TCDD peut induire le cancer simplement par favorisation de la tumorisation des cellules préalablement exposées à des substances initiatrices présentes dans leur environnement, ou génétiquement sujettes à l'initiation spontanée de tumeurs. En plus, les activités biologiques de la TCDD, dont ses propriétés cytotoxiques et promotrices de tumeurs, dépendent entièrement de son affinité envers le récepteur protéique cytosolique *Ah* présent dans une variété de cellules de mammifères. Comme le nombre de complexes récepteur-TCDD dépend de la concentration de TCDD et de l'affinité entre le récepteur et le ligand, la cinétique dose-effet de l'action de promotion de la TCDD devrait présenter un comportement déterministe, ce qui démontrerait l'existence d'un seuil réel pour l'activité de promotion à faibles doses.

À la lumière de ces points de vue mécanistiques, des modifications fondamentales dans la politique de l'EPA pour la réglementation de la TCDD ont été proposées dans un rapport préliminaire paru en 1987 (EPA, 1987), qui recommandait que l'estimation traditionnelle du risque basée sur le modèle multistade de linéarisation soit modifiée par l'inclusion d'une méthode «facteur de sécurité» dans le processus d'évaluation du risque. L'impulsion nécessaire à ce changement provient en partie des difficultés de réglementation qu'a éprouvées l'EPA pour mettre en vigueur la norme rigoureuse d'apport quotidien pour la TCDD à l'aide de la méthode du modèle d'extrapolation, par laquelle une limite d'apport de 0,006 pg/kg/jour suffisait pour produire un risque-repère (DVS) de 10^{-6} pour la vie entière. Cette valeur représentait une norme de réglementation pratiquement irréalisable pour l'exposition de l'homme à la TCDD présente dans le poisson, les effluents toxiques, la terre contaminée et les rejets d'incinérateurs. La limite d'apport de TCDD révisée que l'EPA a proposée dans son rapport préliminaire a été fixée à 0,1 pg/kg/jour, ce qui démontre que le risque estimé de carcinogénicité par suite de l'exposition à la TCDD est 17 fois plus faible lorsque le facteur de sécurité est pris en compte dans l'évaluation du risque.

La limite de dose journalière proposée par l'EPA constitue un hybride particulier représentant le point médian de six différentes évaluations du risque obtenues au moyen des méthodes d'extrapolation vers les faibles doses et DSEO-FS (EPA, 1987; Finkel, 1988; Gough, 1988; Perera, 1988). La justification de l'EPA pour l'inclusion de la méthode DSEO-FS repose sur la possibilité que la TCDD puisse agir comme agent épigénétique favorisant la tumorisation chez l'homme. En conséquence, l'EPA a obtenu une DJA pour la TCDD en ayant recours à une DSEO pour la promotion de tumeurs de 1000 pg/kg/jour (0,001 µg/kg/jour); et l'EPA a choisi un facteur de sécurité soit de 100 soit de 1000 pour produire une valeur de DJA de respectivement 10 et 1 pg/kg/jour. Ces limites de DJA tiennent compte d'une exposition 1670 ou 167 fois plus importante que les limites correspondantes, 0,006 pg/kg/jour/ déterminées du modèle multistade linéarisé.

La valeur de DJA la plus prudente parmi les deux valeurs obtenues par la méthode DSEO-FS (1 pg/kg/jour) est donc «prise pour établir une moyenne» avec la valeur de dose journalière obtenue par la méthode d'extrapolation vers les faibles doses (0,006 pg/kg/jour) pour donner finalement une limite composite de dose journalière de 0,1 pg/kg/jour. Pour une personne de 70 kg, les limites quotidiennes d'exposition selon les méthodes DSEO-FS et d'extrapolation vers les faibles doses sont de 70 pg/jour et de 0,42 pg/jour, pour une moyenne de 7 pg/jour (1 pg/kg/jour × 70 kg). La valeur moyenne de limite d'exposition à la TCDD semble être proche de la moyenne géométrique des limites d'exposition obtenues par les méthodes DSEO-FS et d'extrapolation vers les faibles

doses. La moyenne représente une limite d'exposition 17 fois plus élevée que la limite traditionnelle tirée de la méthode d'extrapolation vers les faibles doses.

La justification qu'a fournie l'EPA pour la limite révisée de la dose journalière est que la valeur obtenue se situe «près du point médian» d'une gamme d'estimations faites au préalable par l'EPA, et d'un groupe de plus faibles estimations effectuées par des organismes de réglementation d'autres pays (dont le Canada), où la méthode du facteur de sécurité est presque la seule à être utilisée (Gough, 1988). L'estimation du risque faite par l'EPA repose donc sur une mise en commun de modèles à seuil et sans seuil, qui sont théoriquement de nature inconciliable. Il n'est pas surprenant que l'EPA, avec ses estimations révisées du risque présenté par la TCDD, ait provoqué de fortes réactions négatives dans certains secteurs: l'agence s'est défendue en déclarant que les changements de méthodologie proposés représentaient simplement une tentative de perfectionnement de l'évaluation du risque (Finkel, 1988).

> *Plutôt, l'EPA déclare publiquement que la dioxine est «un genre» d'agent promoteur de la tumorisation et que sa puissance est «en quelque sorte» inférieure à ce qui a déjà été pensé.*
>
> Finkel (1988)

D'une façon plus pratique, si l'on considère que la dose journalière moyenne de TCDD est d'environ 1 pg/jour aux États-Unis, l'acceptation de la limite d'apport proposée de 7 pg/jour comme norme officielle d'exposition faciliterait grandement l'établissement de normes «raisonnables» d'exposition pour la contamination environnementale par la TCDD (Perera, 1988).

À l'opposé, la U.S. Food and Drug Administration (FDA) a traditionnellement eu recours à une méthode FS simplifiée pour obtenir les niveaux de tolérance des contaminants chimiques dans le poisson de consommation, qui présente généralement des limites d'exposition admissible supérieures aux normes correspondantes de l'EPA (Shu et coll., 1987). Dans le cas de la TCDD, la valeur de DJA est simplement calculée comme le quotient de la DSEO (1000 pg/kg/jour) par un facteur de sécurité modeste de 77, soit 13 pg/kg/jour. La limite prescrite par la FDA est environ 2000 fois plus élevée que la limite classique recommandée par l'EPA, et excède encore de beaucoup la limite révisée de l'EPA. D'après les niveaux estimés de consommation de poisson au 99^{e} percentile, le niveau de tolérance de la FDA pour la TCDD dans le poisson de consommation est fixé à 25 ppb.

2. Méthodologie de l'OME

Dans l'article de revue de Shu et coll. (1987) décrivant l'évaluation du risque de cancer par la TCDD, la démarche prudente d'extrapolation vers les faibles doses préconisée par l'EPA est mise en opposition avec la méthodologie DSEO-FS moins rigoureuse utilisée par le Ministère de l'environnement de l'Ontario. D'après cette méthodologie, l'Ontario et certains pays de l'Europe occidentale considèrent que les estimations du risque de carcinogénicité que présente la TCDD sont significativement plus faibles que les estimations prudentes du risque calculées par l'EPA et d'autres organismes de réglementation aux États-Unis.

La méthodologie de l'OME considère la TCDD comme un promoteur épigénétique à seuil. Par conséquent l'OME utilise un simple modèle FS pour la détermination de la limite de dose journalière, soit: le quotient d'une DSEO de 1000 pg/kg/jour (la même valeur de DSEO que celle choisie par l'EPA pour leur révision) par un facteur de sécurité de 100 résulte en une DJA de 10 pg/kg/jour pour l'homme. Cette valeur est environ 1670 fois plus tolérante que celle calculée par l'EPA, qui est de 0,006 pg/kg/jour.

L'OME a justifié son recours à la méthode DSEO-FS par l'existence d'un seuil théorique à l'activité cancérigène, démontré par l'activité d'agents favorisant la tumorisation chez l'animal, et sur la détermination empirique d'un seuil pour la promotion au moyen d'une valeur de DSEO obtenue de bioessais chroniques. La position de l'OME (selon Shu et coll., 1987) peut se résumer ainsi:

- La DSEO (0,001 µg/kg/jour) déterminée au cours d'études biologiques de carcinogénicité chez le rongeur révèle qu'il existe un niveau de seuil pour la production de tumeur par la TCDD.
- La TCDD a été cotée comme le produit cancérigène le plus puissant au moyen d'un critère utilisant des quantités absolues; il faut donc replacer le tout dans un contexte où les propriétés cancérigènes du produit sont exprimées à des concentrations de seulement deux ou trois ordres de grandeur inférieures à la valeur de la DL_{50}.
- Les études de mutagénicité, d'après une batterie d'épreuves à court terme, révèlent que la TCDD n'est pas un agent mutagène au sens classique du terme. Le manque de preuves permettant de suggérer que la TCDD ou ses métabolites peuvent porter une atteinte physique ou chimique directe à l'ADN appuie également cette conclusion.
- D'après les données de la section ci-haut, il est possible de conclure que les PCDD et les PCDF, surtout la TCDD, peuvent causer des tumeurs chez le rongeur par un mécanisme indirect. Il existe une dose seuil, comme l'indiquent les DSEO tirées d'études à long terme.

Alors, l'évaluation du risque faite par l'OME ne considère la TCDD comme un danger cancérigène qu'à fortes doses proches des niveaux qui induisent une toxicité patente, les effets de promotion se manifestant par l'induction de lésions annonciatrices ou de malignité au niveau des organes cibles. À des doses inférieures à la DSEO, doses auxquelles aucun signe d'une telle toxicité cellulaire ou organique n'est détectable, l'agent est jugé sans danger en raison de la présence d'un véritable seuil de toxicité.

Par contre, l'OME a explicitement rejeté le recours aux modèles d'extrapolation vers les faibles doses pour obtenir des estimations DVS du risque. L'OME mentionne plusieurs points

faibles dans l'évaluation quantitative du risque au moyen de modèles mathématiques qui rendent la méthodologie impropre à l'établissement de normes:

- Il y a un très grand nombre de modèles mathématiques qui produisent des estimations du risque très différentes à partir des mêmes données biologiques.
- Les modèles actuels supposent seulement que le produit chimique agit directement; ils n'incluent pas des modes d'action indirecte (par exemple, les agents favorisant la tumorisation).
- Les modèles actuels extrapolent les probabilités à partir de mesures effectuées entre 10^{-1} et 10^{-2} jusqu'à 10^{-5}, 10^{-6}, ou même 10^{-8}, soit bien en-deçà du domaine de certitude biologique.
- Le recours à ces modèles d'analyse du risque, dans le cas présent, devrait plutôt s'appliquer à définir une gamme possible de doses sûres que de constituer un fondement pour une norme.
- La nature non linéaire, qui est reliée au sexe des données utilisées provenant d'études biologiques chez le rongeur, et la présence d'atteinte hépatique primitive reliée à la dose à des niveaux de traitement causant des modifications néoplasiques au niveau des cellules hépatiques... suggèrent que ces estimations du risque d'induction de cancer peuvent être confondues avec des atteintes tissulaires directes. Le manque de connaissances au sujet du mode d'action de la 2,3,7,8-TCDD exclut également le choix d'un modèle spécifique d'analyse du risque.

Aucun de ces arguments n'est particulièrement nouveau ou avant-gardiste. Cependant, les points qu'ils soulèvent constituent vraiment un exemple pratique du raisonnement que suivent bon nombre d'organismes de réglementation ailleurs qu'aux États-Unis pour choisir la méthode DSEO-FS dans l'établissement des normes relatives aux agents cancérigènes qui ont un comportement de type seuil.

Enfin, l'OME justifie son choix d'un modeste facteur de sécurité de 100 pour la TCDD au moyen des points suivants:

- Le facteur de sécurité de 100 fois représente un moyen pratique de tenir compte des incertitudes inhérentes à l'extrapolation de l'animal vers l'homme. Cette valeur comprend un facteur de 10 fois pour l'extrapolation des données de l'animal à l'homme, car on suppose que l'animal est moins sensible que l'homme, et un second facteur de 10 fois pour tenir compte des différences inter-individuelles de sensibilité.
- Ce facteur incorpore plusieurs considérations prenant en compte l'incertitude inhérente à l'extrapolation de données animales pour l'homme, et particulièrement une marge pour le cas où la sensibilité de l'homme est supérieure à celle de l'espèce soumise à l'épreuve.
- Comme les études de toxicité aiguë et les études à long terme chez l'animal ont été faites, et comme les résultats d'études de mutagénicité à court terme et d'études épidémiologiques chez l'homme sont généralement négatifs, un facteur de sécurité de 100 est recommandé.
- Une DSEO de 0,001 µg/kg/jour pour la 2,3,7,8-TCDD, déterminée par une étude de reproduction couvrant trois générations (Murray et coll., 1979) et une étude oncologique d'une durée de deux ans (Kociba et coll., 1978) (les deux portant sur le rat), est recommandée en guise de base prudente pour la détermination d'une dose journalière maximale admissible de PCDD et de PCDF pour l'homme.

Par conséquent, l'OME a choisi de n'appliquer que le facteur de sécurité traditionnel de 100, sans tenir compte des autres considérations relatives à la nature cancérigène de la TCDD et des incertitudes inhérentes à l'analyse du comportement dose-effet des produits cancérigènes.

En comparaison, plusieurs autres pays qui ont utilisé la méthode DSEO-FS pour la réglementation de la TCDD ont choisi d'introduire des facteurs de sécurité additionnels, ce qui donne des variations notables de la valeur de la DJA (Shu et coll., 1987; Gough, 1988). Aux Pays-Bas, on inclut un facteur de sécurité additionnel de 2,5 pour tenir compte de la possibilité que des effets marginaux non observés puissent se produire au niveau du foie des rates, de sorte que la DSEO apparente serait en fait une DMEO. Cette correction résulte en un facteur de sécurité global de 250, duquel est tirée une DJA de 4 pg/kg/jour. De la même façon, l'agence fédérale de l'environnement en République fédérale d'Allemagne a spécifié un facteur de sécurité qui va de 100 à 1000 fois, ce qui donne une DJA variant entre un et 10 pg/kg/jour. La raison d'être de cette correction supplémentaire du facteur de sécurité est que la TCDD est un contaminant chimique.

F. Modélisation PBP

Pour l'évaluation du risque au moyen de l'extrapolation vers les faibles doses, il est généralement supposé que la probabilité d'induction de tumeur est directement proportionnelle à la dose administrée du produit à l'étude. Toutefois, la plupart des produits cancérigènes subissent un processus pharmacocinétique important dans l'organisme, y compris le métabolisme par l'intermédiaire de voies concurrentielles de bioactivation et de détoxication. Pour un produit cancérigène, il existe donc souvent une relation complexe entre la dose *administrée* (externe) et la dose *efficace* (interne) d'une substance au niveau des organes cibles critiques. La dose efficace est celle qui déterminera le degré d'induction de tumeur au niveau cellulaire. Beaucoup de processus pharmacocinétiques se trouvent saturés à des doses élevées, surtout s'ils mettent en jeu des étapes métaboliques catalysées par des enzymes. En conséquence, la courbe dose-effet d'un produit cancérigène présentera donc souvent

une cinétique non linéaire proportionnelle à la dose: la fonction mathématique décrivant la relation dose-effet variera selon le niveau de dose (Hoel et coll., 1984).

La modélisation PBP (pharmacocinétique basée sur l'aspect physiologique) est une technique destinée à calculer la dose interne d'un produit chimique. Il s'agit d'appliquer un modèle de simulation par ordinateur pour colliger les renseignements au sujet de la dose administrée, des compartiments physiologiques que le produit chimique occupe, et de ses voies métaboliques spécifiques (Andersen et coll., 1987; Watanabe et coll., 1988). La modélisation PBP élabore une description mathématique du devenir du produit chimique dans l'organisme, et fournit des estimations de la concentration de produit cancérigène activé au niveau de la cible. Cette technique ne formule aucune hypothèse quant aux aspects mécanistiques de la carcinogénicité; elle ne constitue qu'une méthode d'estimation plus précise de la dose interne (Andersen et coll., 1987).

Au départ, il est nécessaire de déterminer quelles sont les espèces cancérigènes actives et quelles sont les voies métaboliques possibles qui mènent à la formation de produits cancérigènes en plusieurs étapes dans la cellule. Les modèles de simulation de processus pharmacocinétiques par ordinateur requièrent des données physiologiques approfondies chez l'animal autant que chez l'homme. Il faut déterminer les paramètres du modèle pharmacocinétique pour les deux espèces afin de quantifier les processus d'absorption et de répartition, de bioactivation et de détoxication, et d'excrétion. De plus, il faut également inclure dans le modèle la cinétique de liaison aux macromolécules de la cellule, et les mécanismes de réparation ou d'arrêt du développement des lésions cellulaires.

Malheureusement, sauf pour quelques monomères et solvants organiques bien caractérisés, les données expérimentales relatives à ces paramètres sont souvent incomplètes, surtout chez l'homme (Perera, 1988). Le problème le plus important qui est associé à la modélisation PBP est que la plupart des valeurs des paramètres pharmacocinétiques chez l'homme sont fondées sur des mesures cliniques effectuées à court terme sur un échantillon restreint de la population, ce qui ne reflète pas exactement la grande variation inter-individuelle d'effets chez l'homme. Il est toutefois possible d'introduire dans un modèle PBP un nombre limité de suppositions et d'approximations provenant d'autres sources de données, sans compromettre sérieusement l'exactitude du modèle. Par exemple, comme les constantes pharmacocinétiques de vitesse sont mal définies chez l'homme, il est possible de les remplacer par des constantes métaboliques de vitesse calculées chez l'animal en y ajoutant des facteurs de correction adéquats. La validité des prévisions fondées sur le modèle peut ensuite être évaluée par comparaison des niveaux prévus et observés de la substance chimique mère et de ses métabolites bioactivés dans le sang et les tissus cibles des animaux soumis aux épreuves (Andersen et coll., 1987).

L'impact de la modélisation PBP sur la réglementation s'est récemment manifesté dans l'évaluation de risque proposée par l'EPA dans le cas de deux solvants organiques courants: le chlorure de méthylène (dichlorométhane, DCM) et le 1,1,1-trichloroéthane (TCA). Dans les deux cas, la méthode PBP pour l'évaluation du risque a produit des estimations du risque qui sont significativement plus faibles que celles obtenues à l'aide des techniques classiques. Étant donné l'usage répandu et l'importance économique de ces substances et d'autres solvants aliphatiques contenant du chlore, la méthode PBP a retenu l'attention des organismes de réglementation pour l'établissement de normes révisées concernant la qualité de l'air et de l'eau.

Le chlorure de méthylène (DCM) est un produit cancérigène potentiel qui, à doses élevées, provoque l'induction de tumeurs au poumon et au foie chez la souris. Pour l'homme exposé à des faibles doses de DCM, la modélisation PBP prévoit que la dose interne de métabolites cancérigènes du DCM au niveau des organes cibles est de 140 à 170 fois plus faible pour l'inhalation ou de 50 à 210 fois plus faible pour l'ingestion d'eau potable que la dose équivalente calculée à l'aide de la méthode d'extrapolation vers les faibles doses (Andersen et coll., 1987). Selon la prémisse non vérifiée affirmant que toutes les espèces ont la même sensibilité à une même dose au niveau du tissu cible, le modèle PBP révèle que les évaluations antérieures du risque déterminées au moyen de bioessais chez le rongeur ont pu grandement surestimer le risque de carcinogénicité pour l'homme exposé à de faibles concentrations de DCM.

L'évaluation du risque au moyen de la méthode PBP pour le DCM ne nécessite l'inclusion d'aucun facteur allométrique d'échelle pour tenir compte des différences de taille d'une espèce à une autre; il n'est pas nécessaire non plus d'avoir recours à un facteur de sécurité inter-espèces, puisque le modèle tient compte lui-même des différences pertinentes entre les paramètres pharmacocinétiques de l'homme et ceux de l'animal. Ainsi, la modélisation PBP peut permettre de réduire ou de retirer le facteur de sécurité inter-espèces; un facteur de sécurité tenant compte de la variabilité interindividuelle chez l'homme est cependant toujours nécessaire. Même si plusieurs hypothèses et incertitudes font partie du modèle, les évaluations du risque faites selon la méthode PBP doivent toujours être considérées plus fiables que les évaluations faites selon les méthodes classiques qui ne tiennent pas compte du tout des processus pharmacocinétiques.

La modélisation PBP a également servi à l'évaluation du risque présenté par le 1,1,1-trichloroéthane (TCA), un solvant qui cause une toxicité systémique par l'intermédiaire de l'induction de la prolifération des peroxysomes hépatiques chez les rongeurs (Watanabe et coll., 1988). Les évaluations de risque classiques fondées sur une série de modèles animaux de type «boîte noire» ont suggéré que le TCA pouvait s'accumuler jusqu'à un niveau dangereux au cours d'une exposition chronique, ce qui représenterait alors un risque significatif pour la santé. Par contre, la modélisation PBP prévoit que pour l'homme consommant quotidiennement deux litres d'eau potable contenant 0,3 ppm de TCA, ce dernier ne s'accumulera

pas dans l'organisme et atteindra plutôt un niveau d'équilibre en quelques jours. La charge de l'organisme en TCA à l'équilibre calculée pour l'organisme humain ne sera que de 4×10^{-3} mg/kg, ce qui ne représente que 2,7 fois la dose quotidienne. Ce niveau représente une charge pour l'organisme qui est 5000 fois inférieure au niveau induisant des effets minimaux chez l'animal. Il semble donc qu'il n'y a pas à craindre de risque significatif pour la santé de la contamination des réserves d'eau par le TCA, à des niveaux ordinaires d'exposition pour l'homme.

Dans l'avenir, la modélisation PBP pourrait servir de plus en plus pour l'évaluation du risque présenté par d'autres produits chimiques qui ont une certaine importance économique, comme le trichloroéthylène et le tétrachloroéthylène, à condition toutefois que les données pharmacocinétiques appropriées soient disponibles. Le recours à cette technique réduira certainement les incertitudes inhérentes à l'extrapolation vers les faibles doses dans le cas des produits cancérigènes, et éliminera aussi le facteur de conversion inter-espèces pour tous les types de produits chimiques.

L'utilisation continuelle de facteurs de sécurité arbitraires constitue un biais rigide qui suppose que l'homme est toujours plus sensible que l'espèce animale étudiée la plus sensible, ce qui n'est pas toujours le cas. Si l'on choisit des facteurs de sécurité en fonction des connaissances au sujet de la sensibilité chez l'homme, il faut tenir compte d'autres facteurs comme la gravité de l'effet toxique, du fait que la substance est un médicament possible ou un produit d'usage industriel, du schéma possible d'exposition, et de la rigueur scientifique globale de la base de données. Enfin, on doit constater à quel point le milieu scientifique s'est servi du facteur de sécurité de 100 fois obtenu à partir de la DSEO chez les espèces les plus sensibles. Cependant, comme notre faculté de saisir la sensibilité inter-espèces et de détecter des effets de plus en plus subtils s'améliore toujours, elle justifie le recours à une méthode plus souple que celle qui était utilisée auparavant.

Watanabe et coll. (1988)

La modélisation PBP peut alors éliminer partiellement le besoin d'inclure des facteurs arbitraires de sécurité pour l'établissement des normes.

5. Méthodes FS appliquées aux tératogènes

A. Généralités

Les produits chimiques qui causent des anomalies du développement lors de la reproduction, par toxicité prénatale pour l'embryon ou le foetus, sont appelés tératogènes ou toxiques du développement. Les effets tératogènes se manifestent par une grande variété d'anomalies du développement, dont les fausses couches, la mort foetale, les malformations structurelles, les déficits fonctionnels, la faiblesse du poids à la naissance, ou l'augmentation de la morbidité et de la mortalité néonatales (IRLG, 1986). En termes d'organisation et d'évaluation des données toxicologiques, la méthodologie d'évaluation du risque a été tirée principalement des modèles déjà élaborés pour l'évaluation du risque présenté par les produits cancérigènes. Toutefois, du point de vue mécanistique, les tératogènes sont généralement considérés comme des toxiques à seuil traitables par la méthode DSEO-FS d'évaluation du risque:

> *Le cadre de travail le plus adéquat pour la toxicité au niveau du développement doit être fondé sur une DSE (dose qui n'entraîne pas d'effet) et un facteur de sécurité prudent, parce que la toxicité au niveau du développement est un phénomène de type seuil. Si une substance quelconque peut nuire à l'embryon, on ne doit pas nécessairement la proscrire.*
>
> Johnson (1989)

Pour le moment, la méthodologie d'évaluation du risque la mieux détaillée dans le cas des tératogènes a été décrite dans les EPA *Guidelines for the Health Assessment of Suspect Developmental Toxicants* de 1986 (Lignes directrices relatives à l'évaluation des tératogènes suspects pour la santé) (EPA, 1986c). Toutefois, les méthodes quantitatives d'évaluation sont moins complètes pour les tératogènes que pour les produits cancérigènes:

> *Le SAB (Science Advisory Board/Conseil consultatif pour les questions scientifiques) Panel soulignait que la toxicologie du développement est particulièrement faible quant à l'évaluation quantitative et recommandait que des efforts supplémentaires soient fournis pour mettre au point d'autres méthodes d'estimations quantitatives du risque dans ce domaine.*
>
> EPA (1986c)

Pour l'évaluation dose-effet des tératogènes, il n'est que rarement possible d'utiliser des données adéquates provenant d'études épidémiologiques menées chez l'homme. Par conséquent, afin de pouvoir établir des normes, il faut habituellement effectuer des épreuves biologiques pour obtenir les renseignements dose-effet nécessaires à l'estimation du risque de tératogénicité.

> *La preuve d'une relation dose-effet constitue un critère important pour l'évaluation de la tératogénicité, même si elle doit reposer sur des données restreintes provenant d'études ordinaires à trois doses.*
>
> EPA (1986c)

Malheureusement, il arrive souvent que les données relatives à la tératogénicité chez l'animal ne suivent pas une relation dose-effet classique pour les effets reliés au développement, en raison du grand nombre d'effets possibles et des effets dus aux risques concurrentiels (par exemple, malformations et mort foetale). Il faut donc avoir recours à une mesure combinée de l'évaluation de la tératogénicité, comme mort foetale et malformations, et il faut aussi interpréter avec grand soin chaque effet associé au développement, pour chaque cas (EPA, 1986c). De plus, comme les mesures numériques servant à quantifier les divers effets ont des propriétés métriques différentes, (données en continu, ponctuelles ou tout ou rien), il est impossible de recommander une méthode biostatistique de préférence à une autre pour tous les types de modélisation dose-effet (IRLG, 1986).

Pour la détermination statistique du risque de tératogénicité aux niveaux d'exposition généralement associés à l'environnement ou au milieu de travail, on peut avoir recours à deux types de méthodes d'évaluation du risque à partir des données dose-effet disponibles d'expériences menées sur l'animal. Ces deux méthodes fondamentales sont : 1) la méthode DSEO-FS et 2) les modèles mathématiques

d'extrapolation. La méthode DSEO-FS présuppose l'existence d'un seuil à l'action tératogène, en-deçà duquel aucun signe de toxicité ne se manifeste au niveau du développement. Les modèles d'extrapolation ne présupposent pas l'existence d'un seuil particulier, mais prévoient qu'un niveau de risque de tératogénicité acceptablement faible existera à une dose assez faible de la substance à l'étude.

B. Méthode DSEO-FS appliquée aux tératogènes

D'après les EPA Guidelines de 1986, la méthodologie dose-effet recommandée pour la détermination des limites maximales d'exposition est la méthode du facteur de sécurité (FS) (EPA, 1986c). Il s'agit de déterminer une DSEO (dose sans effet observé) à partir de la dose expérimentale la plus élevée qui n'entraîne aucune augmentation statistique d'anomalies du développement à l'égard d'un effet prédéfini, après soustraction de toute incidence de fond observée dans les groupes témoins. Ensuite, il s'agit de diviser la valeur de la DSEO par un facteur de sécurité, pour tenir compte des incertitudes dues aux différences entre les espèces et à la variabilité humaine: il en résulte la limite maximale d'exposition à un produit donné.

La politique de l'EPA consiste à utiliser une méthodologie DSEO-FS basée sur les exigences des OECD Guidelines #414 relatives à la détermination de la DSEO (OECD, 1981) à la condition qu'un seuil de toxicité soit supposé, mais non prouvé.

La valeur de la DSEO est définie pour les effets les plus sensibles sur le développement, chez l'espèce la mieux appropriée et/ou la plus sensible (EPA, 1986c). Par ailleurs, les 1986 EPA Guidelines spécifient aussi qu'il n'y a pas d'espèce qui soit considérée comme la plus appropriée: dans chaque cas, il faut choisir l'espèce en fonction de l'étude qui sera entreprise. Dans des épreuves effectuées chez l'animal et portant sur un nombre minimal de niveaux de doses (trois), il n'est pas rare de trouver que même la plus faible dose expérimentale entraîne un effet tératogène significatif. La dose la plus faible est alors définie comme la DMEO (dose minimale à effet observé). Des niveaux maximaux d'exposition peuvent être calculés en divisant la DMEO par un nombre comportant un facteur de sécurité additionnel déterminé selon le jugement scientifique. Il est habituel dans la pratique en toxicologie de fixer à 10 la valeur par défaut du facteur de sécurité additionnel pour le calcul de la DMEO.

Même si l'ordre de grandeur du facteur de sécurité est critique au moment de la détermination d'une norme d'exposition, les 1986 EPA Guidelines définissent le niveau d'exposition admissible d'une façon très vague, comme étant: «le niveau calculé d'exposition que l'on croit ne pas être susceptible de causer un effet toxique au niveau du développement de l'homme» (EPA, 1986c).

De cette manière, les 1986 EPA Guidelines ne spécifient aucune valeur prédéterminée pour le facteur de sécurité (facteur d'incertitude). Même si en toxicologie, il est d'usage courant de fixer à 100 la valeur du facteur de sécurité pour les toxiques aigus et chroniques chez l'adulte, l'EPA ne donne aucun critère quantitatif pour le calcul des facteurs de sécurité associés aux tératogènes. L'EPA affirme plutôt que les facteurs de sécurité associés aux tératogènes sont de nature variable, et qu'ils devraient être déterminés un à un, selon un «jugement scientifique». Les EPA Guidelines fournissent toutefois plusieurs critères qualitatifs pour l'évaluation des valeurs de facteur de sécurité:

- *disponibilité d'une DSEO plutôt que d'une DMEO*
  un facteur de sécurité additionnel est nécessaire avec une DMEO
- *pente de la courbe dose-effet*
  une pente peu prononcée produit par extrapolation un risque plus élevé à faible dose, et nécessite l'utilisation d'un plus grand nombre d'animaux à l'étude pour la détermination de la DSEO
- *nature des différences inter-espèces*
  la comparabilité et la cohérence des effets tératogènes observés chez les animaux à l'étude auront une influence sur le facteur de sécurité
- *types d'effets observés au niveau du développement*
  des effets plus graves nécessitent le recours à un facteur de sécurité plus élevé
- *rapport de doses de toxicité maternelle/foetale (rapport A/D)*
  des effets spécifiquement foetotoxiques nécessitent le recours à un facteur de sécurité plus élevé

En général, les EPA Guidelines préconisent également le choix d'un facteur de sécurité global tenant compte de l'importance des preuves de tératogénicité, de la fidélité de la relation dose-effet, du degré d'exposition humaine, et de «tout autre facteur pouvant affecter la qualité et la précision de l'évaluation» (EPA, 1986c). L'EPA souligne à plusieurs reprises le rôle du «jugement d'expert» dans l'application de ces critères à l'établissement des facteurs de sécurité.

Une méthodologie mise au point par la U.S. Food and Drug Administration (U.S. FDA) pour les évaluations de la sécurité des médicaments (Frankos, 1985) traite d'une manière approfondie le problème qui consiste à déterminer les valeurs appropriées pour les facteurs de sécurité. Cette méthode comporte une évaluation intégrée de la toxicité maternelle et de la tératogénicité à l'intérieur de l'évaluation du risque; les substances sont ainsi classées selon: 1) la «*spécificité*» des effets tératogènes «*au niveau du développement*» des animaux à l'étude, et 2) la «*sélectivité embryotoxique*» de la relation dose-effet de tératogénicité.

La «*spécificité au niveau du développement*» indique si les effets observés en relation avec le développement sont spécifiques ou non (Frankos, 1985; IRLG, 1986). Les effets «spécifiques» (Tableau 5.1) consistent en des anomalies structurelles et des effets létaux pour l'embryon (Type I); ils sont considérés comme plus graves et assurément plus tératogènes. Des effets «non spécifiques» comme les retards de développement du foetus, la faiblesse du poids à la naissance, et la diminution de la survie et de la croissance post-natales

(Type II) sont considérés comme moins graves et moins indicateurs de tératogénicité. Qualitativement, les effets du Type I devraient avoir plus d'importance que ceux du Type II; ce critère devrait donc affecter la valeur du facteur de sécurité.

Tableau 5.1
Effets de tératogénicité

Type I	
↓du nombre de naissances vivantes (taille de la portée)	↓du nombre de foetus vivants (taille de la portée)
↑du nombre de mortinaissances	↑du nombre de résorptions
	↑du nombre de foetus avec des malformations, des changements structurels
Type II	
↓du poids à la naissance	
↓de la survie post-natale	↑du nombre de foetus présentant des retards de développement (squelette, tissus mous)
↓de la croissance post-natale, de la capacité reproductive	

Source: Frankos (1985).

La «*sélectivité embryotoxique*», souvent exprimée comme étant le *rapport A/D*, désigne la relation entre la dose la plus faible (A) qui produit des signes évidents de toxicité chronique chez la mère, et la dose la plus faible (D) induisant des signes de tératogénicité chez l'embryon. Les substances dont le rapport A/D est petit sont considérées comme des agents «coefficaces» qui produisent des effets toxiques dans la même gamme de doses chez la mère que chez le foetus. Ce mode d'action est souvent attribuable à un effet secondaire sur le développement foetal causé par une homéostase maternelle altérée (Johnson, 1989). En général, les agents dont le A/D est <1 ont un indice faible de danger et ne présentent pas de sélectivité embryotoxique. Par contre, les agents dont le A/D est >1 peuvent présenter un danger réel pour le développement.

Les effets toxiques pour la mère sont généralement non spécifiques et facilement détectables dans les populations humaines; en conséquence, les substances dont le rapport A/D est petit sont considérées relativement moins dangereuses. Par contre, les substances dont le rapport A/D est grand présentent des effets toxiques à des doses de beaucoup inférieures aux doses toxiques pour la mère. Ces substances sont considérées comme étant des toxiques moins dangereux pour le développement, parce que les limites d'exposition adéquates pour protéger les adultes des effets nocifs pour la santé seraient insuffisantes pour protéger le foetus exposé aux effets tératogènes. Par exemple, la thalidomide possède un rapport A/D élevé, supérieur à 30, alors que la vitamine A possède un rapport A/D d'environ deux. Les deux substances sont tératogènes, mais seules les doses de vitamine A proches du niveau de toxicité chez l'adulte produiront des effets foetotoxiques, alors que des doses de thalidomide beaucoup plus faibles sont extrêmement tératogènes.

Une variété de mesures dose-effet définissant précisément le rapport A/D ont été proposées; cependant aucune d'entre elles n'a été acceptée de façon répandue jusqu'à maintenant (Bass et coll., 1985; Fabro, 1985; Johnson, 1989). Une définition quantitative plus explicite du rapport A/D a été suggérée comme «*Indice de tératogénicité relative*»; elle représente le rapport DL_{01} (dose létale) maternelle/DT_{05} (dose tératogène) pour l'embryon (Fabro, 1985). Cette mesure peut être intégrée en tant que critère additionnel dans la détermination des facteurs de sécurité pour les tératogènes; elle n'a cependant pas encore été acceptée partout et a fait l'objet de critiques pour des motifs méthodologiques (Johnson, 1989).

La méthode de la FDA relative à la détermination des facteurs de sécurité pour les tératogènes spécifie des valeurs variant entre 100 et 1000, selon principalement une matrice de la «spécificité du développement» (Tableau 5.2). Cette procédure permet de vérifier qu'un type I ou II de tératogénicité a bien été détecté, au moyen: 1) d'une simple épreuve de tératogénicité, ou bien 2) d'une épreuve couvrant plusieurs (deux) générations. Si les deux épreuves révèlent la présence d'effets de type I, alors il faut attribuer un facteur de sécurité de 1000. Si les épreuves ne révèlent que la présence d'effets de type II, ou une combinaison effets de type II/pas d'effet, alors le facteur de sécurité est de 100. Si les épreuves révèlent la présence d'effets de type I et d'autres effets au niveau du développement, il faut alors déterminer, à l'aide de critères qualitatifs additionnels semblables à ceux suggérés par les EPA Guidelines, une valeur intermédiaire de facteur de sécurité située entre 100 et 1000.

Tableau 5.2
Facteurs de sécurité appliqués aux résultats d'épreuves de tératogénicité

	Résultats d'épreuves menées sur plusieurs générations		
Épreuves tératologiques	Type I	Type II	Sans effet
Type I	Type I (1000×)	Type I ( )*	Type I ( )*
Type II	Type I ( )*	Type II (100×)	Type II (100×)
Sans effet	Type I ( )*	Type II (100×)	Sans effet (100×)

*Un facteur de sécurité approprié sera déterminé après examen de renseignements additionnels au sujet du métabolisme, de la pharmacocinétique, de la toxicité naturelle, des mécanismes, etc.

Source: Frankos (1985).

La méthode de la FDA peut produire des valeurs de facteurs de sécurité plus systématiques que la méthode de l'EPA, même si elle sert surtout dans le cas des tératogènes bien caractérisés dont les ensembles de données proviennent à la fois d'observations faites au niveau du développement lors d'expériences simples et lors d'expériences couvrant plusieurs générations.

C. Méthodes mathématiques de modélisation appliquées aux tératogènes

L'autre façon d'aborder la question de l'évaluation du risque est l'utilisation de modèles mathématiques d'extrapolation. Toutefois elle n'est pas bien acceptée, selon les EPA Guidelines de 1986:

> *Pour le moment, il n'existe aucun modèle mathématique qui soit généralement accepté pour l'estimation des effets tératogènes en-deçà de la gamme de doses appliquée. Ceci est dû au manque de connaissances permettant de comprendre les mécanismes biologiques de la tératogénicité, les différences interindividuelles et inter-espèces quant à la nature des observations liées au développement, l'influence des effets maternels sur la courbe dose-effet, et la présence ou l'absence de seuil sous lequel un agent ne peut pas produire d'effets.*
>
> EPA (1986c)

Les modèles mathématiques décrivant en détail les processus tératogènes n'ont pas été adéquatement élaborés; des procédures mathématiques générales d'extrapolation peuvent toutefois être imaginées, comme celles proposées par Gaylor (1983, 1988), Crump (1984) et Dourson (1986), pour l'analyse dose-effet des toxiques systémiques. De telles méthodes mathématiques pourraient convenir à l'analyse dose-effet des données de compte binomiales (ou rapports de données de compte), comme la taille de la portée, ou le nombre de malformations par portée. La méthode de Gaylor est un peu analogue à la méthode d'extrapolation vers les faibles doses utilisée pour l'estimation du risque présenté par les produits cancérigènes, mais dans le contexte présent, elle n'exclut pas l'existence d'un seuil réel.

L'avantage principal des modèles mathématiques pour l'estimation du risque est qu'ils offrent la possibilité de calculer de façon plus objective le risque à faibles doses, et qu'ils éliminent la nécessité d'émettre des hypothèses individuelles (cas par cas) quant à la grandeur du facteur de sécurité. Ces méthodes fournissent une estimation de la limite supérieure du risque présenté par de faibles doses de tératogènes; il s'agit donc de méthodes dont les hypothèses sont prudentes, et qui peuvent ainsi donner des valeurs quantitatives comparables à celles obtenues par la méthode FS (Gaylor, 1983). La comparaison de la grandeur relative du risque calculée selon la méthode DSEO-FS et de celle calculée selon la procédure mathématique d'extrapolation favoriserait aussi la validation des estimations du risque faites selon la méthodologie DSEO-FS.

D. Extrapolation inter-espèces

Les EPA Guidelines de 1986 n'abordent pas les questions méthodologiques touchant l'extrapolation inter-espèces, de l'animal de laboratoire à l'homme, des données dose-effet. Elles recommandent plutôt d'effectuer l'extrapolation inter-espèces selon chaque cas pris séparément, en fonction des similarités biologiques existant entre l'espèce étudiée et l'homme:

> *La détermination de la grandeur des facteurs de sécurité pour la conversion inter-espèces ne relève pas des mathématiques, mais plutôt de l'interprétation biologique.*
>
> IRLG (1986)

Les preuves de la validité des résultats d'épreuves animales servant à prévoir les effets tératogènes chez l'homme ne sont pas très abondantes, mais il existe un bonne concordance générale entre la toxicité pour l'embryon animal et celle pour l'embryon humain, pour la plupart des substances:

> *Ces comparaisons (animal-homme) montrent presque toujours une concordance entre les effets observés chez l'homme et ceux observés chez au moins une espèce animale soumise aux épreuves... Il existe donc une certaine base pour l'estimation, à partir de données tirées d'études animales, du risque que présente l'exposition pour le développement humain.*
>
> EPA (1986c)

En général, les effets tératogènes chez l'animal et chez l'homme semblent être qualitativement et quantitativement similaires, de sorte que l'embryon humain ne semble pas significativement plus sensible aux effets tératogènes que l'embryon animal (Johnson, 1989). Par conséquent, la grandeur du facteur de sécurité nécessaire pour tenir compte des différences de sensibilité aux tératogènes d'une espèce à une autre ne devrait pas dépasser celle concernant les toxiques systémiques, même si l'imprévisibilité des effets tératogènes d'une espèce à une autre entraînera quelquefois un risque plus grand chez l'homme (par exemple, la thalidomide) que celui qui avait été prévu lors des expériences animales.

E. Problèmes spéciaux reliés à la détermination des FS pour les tératogènes

Les méthodes d'évaluation du risque associé aux dangers de tératogénicité comprend plusieurs caractéristiques spéciales dont il n'est généralement pas question dans les évaluations classiques du risque. De simples mesures du niveau et de la durée de l'exposition de la mère à un tératogène ne constituent pas toujours un élément de prévision fiable quant à l'importance de l'effet tératogène au niveau des sites cibles à l'intérieur de l'embryon en développement, puisque la relation dose-effet est étroitement reliée au stade de développement. Ainsi, une simple exposition de courte durée peut suffire à entraîner des effets délétères pour le développement au cours d'une période critique de l'organogénèse; les doses tératogènes sont souvent de beaucoup inférieures aux doses de toxicité chronique chez la mère (OMS, 1984).

De plus, le groupe spécifique à risque élevé quant à l'exposition aux tératogènes ne comprend pas seulement les femmes visiblement enceintes, mais toutes les femmes fertiles

en âge de reproduction, puisqu'il arrive souvent que des femmes ne se sachent pas encore enceintes au moment des premiers stades de l'embryogénèse. En fait, la période pendant laquelle l'embryon est le plus sensible aux effets des tératogènes couvre les quatre premières semaines de grossesse, au moment où, dans beaucoup de cas, la femme ne se sait pas encore enceinte. Par conséquent, les mesures de précaution visant à retirer des lieux suspects des femmes visiblement enceintes qui subiraient une exposition importante à des tératogènes n'apporte aucune protection pour les embryons aux premiers stades de développement. Il faut donc examiner des critères sélectifs d'exposition des femmes fertiles à des substances potentiellement tératogènes, surtout pour établir des normes pour le milieu de travail. De telles considérations soulèvent de sérieux problèmes quant à l'égalité en matière de travail et les dispositions des lois du travail relatives à la discrimination basée sur le sexe.

F. Résumé

D'après les EPA Guidelines de 1986, lorsqu'il s'agit d'estimer le risque, la première méthode à laquelle il faut recourir pour déterminer les limites admissibles d'exposition à des tératogènes devrait être la DSEO-FS. Il serait toutefois possible de présenter une limite maximale d'exposition, déterminée d'après une méthode mathématique d'extrapolation, dans le cas des effets susceptibles d'être traités de cette façon, par exemple, ceux dont les données sont liées à des données de compte ou de rapport. Le choix de l'une ou l'autre méthode doit être fait selon un jugement scientifique, en fonction des propriétés biostatistiques des données dose-effet et des effets associés au développement qui ont été mis en évidence au cours des expériences effectuées chez l'animal.

Les DSEO estimées et les courbes mathématiques d'extrapolation ont tendance à présenter de grandes marges d'erreur, en raison de la petite taille de l'échantillon utilisé pour la détermination des points individuels de la courbe dose-effet, et de la grande variabilité d'effets tératogènes chez les animaux soumis aux épreuves. Ainsi, le risque de tératogénicité devrait être estimé à la limite supérieure de l'intervalle de confiance à 95% pour les deux méthodes. La meilleure estimation de la DSEO, ou la courbe la mieux ajustée en fonction de l'extrapolation pourrait aussi être présentée à des fins de comparaison, mais ces valeurs ne devraient pas servir à la détermination des limites d'exposition.

Dans le cas de la méthode DSEO-FS, le choix d'une valeur appropriée pour le facteur de sécurité représente une source importante d'incertitude pour l'estimation des limites d'exposition aux tératogènes. Dans l'évaluation du risque, il faudrait spécifier la valeur du facteur de sécurité utilisé, et présenter une discussion des hypothèses émises et de l'exactitude des estimations qui ont servi à déterminer cette valeur. Il est d'usage courant de choisir 100 comme valeur minimale, mais il est également possible de choisir un facteur de sécurité plus grand, en fonction de critères spéciaux dont il a été question plus haut.

Bibliographie

ACGIH (1986). American Conference of Government Industrial Hygienists. *Identification and Classification of Carcinogens: Procedures of the Chemical Substances Threshold Limit Value Committee*, ACGIH. Am J Public Health 76: 1232-1235.

Aldridge WN (1986). *The Biological Basis and Measurement of Thresholds*. Ann Rev Pharmacol Toxicol 26: 39-58.

Andersen ME, Clewell HJ, Gargas ML, Smith FA, Reitz RH (1987). *Physiologically Based Pharmacokinetics and the Risk Assessment Process for Methylene Chloride*. Toxicol Appl Pharmacol 87: 185-205.

Anderson EL (1985). *Quantitative Approaches in Use in the United States to Assess Cancer Risk*, in: Vouk VB et al, Eds., Methods for Estimating Risk for Chemical Injury: Human and Non-Human Biota and Ecosystems (SCOPE 26). John Wiley and Sons, Toronto, pp 405-436.

Ashford NA, Ryan CW, Caldart CC (1983). *Law and Science Policy in Federal Regulation of Formaldehyde*. Science 222: 894-900.

Barnes DG et Dourson ML (1988). *Reference Dose (RfD): Description and Use in Health Risk Assessments* Regul Toxicol Pharmacol 8: 471-486.

Bass R, Neubert D, Stotzer H, Bochert G (1985). *Quantitative Dose-Response Models in Prenatal Toxicology*, in Vouk VB et coll., Eds., Methods for Estimating Risk of Chemical Injury: Human and Non-Human Biota and Ecosystems (SCOPE 26). John Wiley and Sons, Toronto, pp 437-456.

Bigwood EJ (1973). *The Acceptable Daily Intake of Food Additives*. CRC Crit Rev Toxicol (June): 41-93.

Calabrese EJ (1978a). *Methodological Approaches to Deriving Environmental and Occupational Health Standards*. John Wiley and Sons, New York.

Calabrese EJ (1987b). *Pollutants and High Risk Groups*. John Wiley and Sons, New York.

Calabrese EJ (1983). *Principles of Animal Extrapolation*. John Wiley and Sons, New York.

Calabrese EJ (1985). *Uncertainty Factors and Interindividual Variation*. Regul Toxicol Pharmacol 5: 190-196.

Calabrese EJ (1987). *Animal Extrapolation: A Look Inside the Toxicologist's Black Box*. Environ Sci Technol 21:618-623.

Cotruvo JA (1988). *Drinking Water Standards and Risk Assessment*. Regul Toxicol Pharmacol 8: 288-299.

Crump KS (1984). *A New Method For Determining Allowable Daily Intakes*. Fundam Appl Toxicol 4: 854-871.

Crump KS (1986). *Letter*. Fundam Appl Toxicol 6: 183-184.

Dourson ML (1986). *New Approaches in the Derivation of Acceptable Daily Intake (ADI)*. Comments Toxicology 1:35-48.

Dourson ML, Stara JF (1983). *Regulatory History and Experimental Support of Uncertainty (Safety) Factors*. Regul Toxicol Pharmacol 3: 224-238.

Dourson ML, Hertzberg RC, Hartung R, Blackburn K (1985). *Novel Method for the Estimation of the Acceptable Daily Intake*. Toxicol Ind Health 1:23-41.

Dourson ML, Hertzberg RC, Stara JF (1986). *Letter*. Fundam Appl Toxicol 6: 182-183.

EPA (1980). U.S. Environmental Protection Agency. *Guidelines and Methodology Used in the Preparation of Health Effects Assessment Chapters of the Consent Decree Water Quality Criteria*. Federal Register 45: 79347-79357.

EPA (1985). U.S. Environmental Protection Agency. *National Primary Drinking Water Regulations; Synthetic Organic Chemicals, Inorganic Chemicals, and Microorganisms*. Federal Register 50: 46936-47025.

EPA (1986a). U.S. Environmental Protection Agency. *Guidelines for the Health Risk Assessment of Chemical Mixtures*. Federal Register 51: 34014-34025.

EPA (1986b). U.S. Environmental Protection Agency. *Guidelines for Carcinogen Risk Assessment*. Federal Register 51: 33992-34003.

EPA (1986c). U.S. Environmental Protection Agency. *Guidelines for the Health Assessment of Suspect Developmental Toxicants*. Federal Register 51: 34028-34040.

EPA (1987). *A Cancer Risk Specified Dose Estimate for 2,3,7,8-TCDD: Review Draft*. Environmental Protection Agency, Washington DC.

Fabro S (1985). *On Predicting Environmentally-Induced Human Reproductive Hazards: An Overview and Historical Perspective*. Fundam Appl Toxicol 5: 609-614.

Finkel AM (1988). *Dioxin: Are We Safer Now Than We Were Before?* Risk Analysis 8: 161-165.

Frankos VH (1985). *FDA Perspectives on the Use of Teratology Data for Human Risk Assessment*. Fundam Appl Toxicol 5: 615-625.

Gaylor DW (1983). *The Use of Safety Factors for Controlling Health Risks*. J Toxicol Environ Health 11: 329-336.

Gaylor DW (1988). *Applicability of Cancer Risk Assessment Techniques to Other Effects*. Toxicol Ind Health 4: 453-459.

Gaylor DW, Kodell RL (1980). *Linear Interpolation Algorithm for Low Dose Risk Assessment of Toxic Substances*. J Environ Pathol Toxicol 4: 305-312.

Gough M (1988). *Science Policy Choices and the Estimation of Cancer Risk Associated with Exposure to TCDD*. Risk Analysis 8: 337-342.

Hartung R, Durkin PR (1986). *Ranking the Severity of Toxic Effects: Potential Applications to Risk Assessment*. Comments Toxicology 1:49-63.

Hattis D (1987). *The Value of Molecular Epidemiology in Quantitative Health Risk Assessment*, in: Draggan S et al., Eds., Environmental Impacts on Human Health. Preager, New York.

Hattis D, Erdreich L, Ballew M (1987). *Human Variability in Susceptibility to Toxic Chemicals — A Preliminary Analysis of Pharmacokinetic Data from Normal Volunteers*. Risk Analysis 7: 415-426.

Hoel DG, Kaplan NL, Anderson MW (1984). *Implication of Non-linear Kinetics on Risk Estimation in Carcinogenesis*. Science 219: 1032-1036.

Hutt PB (1985). *Use of Quantitative Risk Assessment in Regulatory Decisionmaking under Federal Health and Safety Statutes*, in: DG Hoel et coll., Eds., Risk Quantitation and Regulatory Policy (Banbury Report 19). Cold Spring Harbor Laboratory, New York.

IRLG (1979). Interagency Regulatory Liaison Group, Work Group on Risk Assessment. *Scientific Bases for the Identification of Potential Carcinogens and Estimation of Risks*. Federal Register 44: 39858-39879. [également publié dans: Annual Review of Public Health 1: 345-393 (1980)].

IRLG (1986). Interagency Regulatory Liaison Group (Kimmel CA, Kimmel GL, and Frankos VH, Eds.). *Interagency Regulatory Liaison Group Workshop on Reproductive Toxicity Risk Assessment*. Environ Health Perspect 66: 193-221.

Johnson EM (1989). *A Case Study of Developmental Toxicity Risk Estimation Based on Animal Data: The Drug Bendectin*, in: DJ Paustenbach, Ed., The Risk Assessment of Environmental and Human Health Hazards: A Textbook of Case Studies. John Wiley and Sons, New York.

Kimmel CA, Gaylor DW (1988). *Issues in Qualitative and Quantitative Risk Analysis for Developmental Toxicology*. Risk Analysis 8:15-20.

Kociba RJ, Keyes DG, Beyer JE, et coll. (1978). *Results of a Two-year Chronic Toxicity and Oncogenicity Study of 2,3,7,8-TCDD in rats*. Toxicol Appl Pharmacol 46: 279-303.

Krewski D, Brown C, Murdoch D (1984). *Determining "Safe" Levels of Exposure: Safety Factors or Mathematical Models?* Fundam Appl Toxicol 4: S383-S394.

Kroes R, Feron VJ (1984). *General Toxicity Testing: Sense and Nonsense, Science and Policy*. Fundam Appl Toxicol 4: S298-S308.

Lehman AJ, Fitzhugh OG (1954). *100-fold Margin of Safety*. Assoc Food Drug Off U.S. Q Bull 18: 33-35.

Lu FC (1983). *Toxicological Evaluations of Carcinogens and Non-Carcinogens: Pros and Cons of Different Approaches*. Regul Toxicol Pharmacol 3:121-132.

Lu FC (1985). *Safety Assessments of Chemicals with Thresholded Effects*. Regul Toxicol Pharmacol 5: 460-464.

Lu FC (1988). *Acceptable Daily Intake: Inception, Evolution, and Application*. Regul Toxicol Pharmacol 8: 45-60.

Mantel N (1975). *The Assessment of Carcinogenicity*. Cancer Research 35:865-876.

Marshall E (1982). *EPA's High Risk Carcinogen Policy*. Science 218: 975-982.

Munro IC, Krewski DR (1981). *Risk Assessment and Regulatory Decision-Making*. Fd Cosmet Toxicol 19: 549-560.

NAS (1977). *Drinking Water and Health*. National Research Council, U.S. National Academy of Sciences. Washington DC.

Nelson N (1984). *Use of Toxicity Test Data in the Estimation of Risks to Human Health*, in: Rodricks JV and Tardiff RG, Eds., Assessment and Management of Chemical Risks (ACS Symposium Series 239). American Chemical Society, Washington.

OECD (1981). Organization for Economic Co-Operation and Development. *OECD Guidelines for Testing of Chemicals #414 -Teratogenicity*.

OSTP (1986). Office of Science and Technology Policy, U.S. Interagency Staff Group on Carcinogens (Hart RW, Chairman). *Chemical Carcinogenesis: A Review of the Science and its Associated Principles*. Environ Health Perspect 67: 201-282.

Perera FP (1988). *EPA Cancer Risk Assessment (letter)*. Science 239: 1227.

Perera FP, Petito C (1982). *Formaldehyde: A Question of Cancer Policy*. Science 216: 1285-1291.

Preuss PW, Erlich AM (1987). *Environmental Protection Agency's Risk Assessment Guidelines. Health of Non-Criteria Pollutants*. J Air Pollution Control Assoc 37: 783-797.

Rodricks JV, Tardiff RG (1984). *Conceptual Basis for Risk Assessment*, in: Rodricks JV and Tardiff RG, Eds., Assessment and Management of Chemical Risks (ACS Symposium Series 239). American Chemical Society, Washington.

Rushefsky ME (1985). *Assuming the Conclusions: Risk Assessment in the Development of Cancer Policy*. Politics and the Life Sciences 4: 3144.

Sharratt M (1976). *Uncertainties Associated with the Evaluation of the Health Hazards of Environmental Chemicals from Toxicological Data*, in: Hunter WJ and Smeets JG, Eds., The Evaluation of Toxicological Data for the Protection of the Environment, Pergamon Press, Toronto.

Shu HP, Paustenbach DJ, Murray FJ (1987). *A Critical Evaluation of the Use of Mutagenesis, Carcinogenesis and Tumor Promotion Data in a Cancer Risk Assessment of 2,3,7,8-Tetrachlorodibenzo-p-Dioxin*. Regul Toxicol Pharmacol 7: 57-88.

Somers E (1981). *Environmental Risk Management in Canada*. Regul Toxicol Pharmacol 3: 75-81.

Somers E (1984). *Risk Estimation for Environmental Chemicals as a Basis for Decision-making*. Regul Toxicol Pharmacol 4: 99-106.

Stara JF, Bruins RJF, Dourson ML, Erdreich LS, Hertzberg RC, Durkin PR, Pepelko WE (1987). *Risk Assessment is a Developing Science: Approaches to Improve Evaluation of Single Chemicals and Chemical Mixtures*, in: VB Vouk et al, Eds., Methods for Assessing the Effects of Mixtures of Chemicals (SCOPE 30). John Wiley and Sons, Toronto.

Vettorazzi G (1976). *Safety Factors and their Application in the Toxological Evaluation*, in: WJ Hunter and JG Smeets, Eds., The Evaluation of Toxicological Data for the Protection of Public Health. Pergamon Press, Oxford.

Vocci F, Farber T (1988). *Extrapolation of Animal Toxicity Data to Man*. Regul Toxicol Pharmacol 8: 389-398.

Watanabe PG, Schumann AM, Reitz RH (1988). *Toxicokinetics in the Evaluation of Toxicity Data*. Regul Toxicol Pharmacol 8: 408-413.

Weil CS (1972). *Statistics versus Safety Factors and Scientific Judgment in the Evaluation of Safety Factors for Man*. Toxicol Appl Pharmacol 21: 454-463.

Weil CS, McCollister DD (1963). *Relationship between short- and long-term Feeding Studies in Designing an Effective Toxicity Test*. Agric Food Chem 11: 486-491.

Weisberger JH et Williams GM (1981). *Carcinogen Testing: Current Problems and New Approaches*. Science 214: 401-407.

WHO (1984). World Health Organization. *Principles for Evaluating Health Risks to Progeny Associated with Exposure to Chemicals During Pregnancy*. Environmental Health Criteria 30. World Health Organization (Organisation mondiale de la santé), Geneva.

WHO (1987). World Health Organization. *Principles for the Safety Assessment of Food Additives and Contaminants in Food*. Environmental Health Criteria 70. World Health Organization (Organisation mondiale de la santé), Geneva.